```
1 AVTISVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92.1%;
93.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            45824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 353; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24
400
286
400
400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            <1
25
25
287
383
400 /
   ANCHOR
1151.5
1145.5
1145.5
1145.5
1147.5
1134
1134
1134
1136
1063.5
1063.5
1063.5
1064.5
1048.5
1047.5
1047.5
1047.5
1047.5
1047.5
1047.5
1047.5
1047.5
1047.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
CHAIN
TRANSMEM
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             666800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         003999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
Q03999
   HID DOLL OF THE LEAD OF THE LAND OF THE LA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                       (without alignments)
711.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              plasmodium
plasmodium
plasmodium
plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          lasmodium
lasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            plasmodium
plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     plasmodium
plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q03999 plasmodium
Q25966 plasmodium
Q25976 plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           lasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                 ; Search time 70.07 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
                                                                                                                                                                                                                                            1 AVTISVIDNILSKIENEYEV......SNFLGISFLLILMLILYSFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q25984
Q25972
Q25972
Q25973
Q25973
Q25973
Q25974
Q25975
Q25977
Q25977
Q25977
Q25977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q9tzt4
Q25983
                   GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                              425026 seqs, 132305027 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Maximum Match 100%
Listing first 45 summaries
                                                                                             protein search, using sw model
                                                                                                                               August 8, 2001, 12:33:46
                                                                                                                                                                                                                                                                                      BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_rodent:*
sp_unclassified:*
sp_vertebrate:*
sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         003999
025966
025976
025973
025973
025973
025974
025974
025974
025977
025977
025977
025977
025977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q9TZT4
Q25983
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sp_invertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sp_organelle:*
                                                                                                                                                                                                                                                                                                                                                                                                                        Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Post-processing: Minimum Match 0%
                                                                                                                                                                                                             US-09-500-376-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_fungi:*
sp_human:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sp_mammal:*
sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_phage:*
sp_plant:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SPTREMBL_16:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               992.1
992.1
992.1
992.1
992.1
992.1
988.8
988.8
988.8
988.8
988.8
988.8
988.8
988.8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10:
                                                                                                                                                                                                                                                                                                                                                                                    Total number of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1824.5
1824.5
1824.5
1824.5
1824.5
1821.5
1821.5
1821.5
1821.5
1821.5
183.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1164.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
1165.5
116
                                                                                                                                                                                                                                                                                    Scoring table:
                                                                                                                                                                                                                              Perfect score:
                                                                                           OM protein
                                                                                                                                                                                                                                              Sequence:
                                                                                                                                                                                                                                                                                                                                                Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Database
                                                                                                                                 ë
G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Result
                                                                                                                                 Run
```

```
025970 plasmodium
025980 plasmodium
025982 plasmodium
025978 plasmodium
025972 plasmodium
025972 plasmodium
025722 plasmodium
025721 plasmodium
025721 plasmodium
025721 plasmodium
025722 plasmodium
025728 plasmodium
025718 plasmodium
025718 plasmodium
025718 plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
Blackman M.J. Ling I.T., Nicholls S.C., Holder A.A.;
Submitted (XXX-1991) to the EMBL/GenBank/DDBJ databases.
-i- SUBCELLULAR LOCATION: COULD BE ATTACHED TO THE MEMBRANE BY A GPI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               83K MEROZOITE SURFACE ANTIGEN.
42K MEROZOITE SURFACE ANTIGEN.
19K MEROZOITE SURFACE ANTIGEN.
MEMBRANE ANCHOR.
537F075058626AC2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-VUN-2000 (TrEMBLrel. 14, Last annotation update)
MEROZOITE SURFACE ANTIGEN PRECURSOR I (PMMSA) (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                400 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ALIGNMENTS
   025970
025980
025980
025968
025973
025973
025724
025727
043995
025727
043995
025727
043995
025727
043995
025727
043995
025727
043995
025727
025727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                   043996
025728
025718
09UB87
09NCN2
```

1;

Gaps

1;

Score 1824.5; DB 5; Length 400; Pred. No. 1.5e-87; 6; Mismatches 17; Indels 1;

M.

a

25

셤 ò a ò g ò g οy g à q ò d

85 121

61

145

181 205 241 265 384

RESULT Q25966

ACCOUNT OF THE STANT OF THE STA

361

301

```
LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                      VPENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                                  539
                                                                                                                                                                                                                                                                                                                                                                                                                            Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13362; BAA02623.1; -.
InterPro; IPR000561; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PROTEIN_KINASE_TYR;
                                                                                                                                                                                                                                                        Created)
                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        539 AA; 61045 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           92.1%;
93.6%;
                                                                                                                                                                                                                                                       01,
                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Merozoite; EGF-like domain.
                                                                                                                                                   LGISFLLILMLILYSFI 377
                                                                                                                                                              Query Match 92.1
Best Local Similarity 93.6
Matches 353; Conservative
                                                                                                                                                                                                                                                      01-NOV-1996 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001245; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF00008; EGF; 1
PROSITE; PS00109; PRO1
                                                                                                                                                                                                                                                                                                                                                  [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                025976
025976;
                                                                                                                                                                                                            m
                                                                                                                                                                         523
                                                          241
              181
                                                                                                                                                   361
                                                                                                      301
                                                                                                                            463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      301
                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                      025976
                                QQ
                                                         ò
                                                                               g
                                                                                                      ò
                                                                                                                            g
                                                                                                                                                                        g
                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
             ò
                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ä
                                                                                180
                                                                                                                            240
                                                                                                                                                                         300
                                                                                                                                                                                              323
                                                                                                                                                                                                                     360
                                             204
                                                                                                                                                  264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV
                                                                                           VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                     LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                    GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal part of the precursor to major merozoite surface proteins (MSP1) of Plasmodium falciparum field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          539;
                                                                                                                                                                                                                                                                                                                                                                                                                            Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 1824.5; DB 5; Length
Pred. No. 2e-87;
6; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7B7DE90C1D0ACDE7 CRC64;
                                                                                                                                                                                                                                                                                                                                                              01-NOV-1996 (TYEMBLEEL. 01, Created)
01-NOV-1996 (TYEMBLEEL. 01, Last sequence update)
01-JUN-2000 (TYEMBLEEL. 14, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
MSP1.
                                                                                                                                                                                                                                                                                                                                                 Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13357; BAA02618.1; -.
InterPro; IPR000561; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EGF; 1.
)9; PROTEIN_KINASE_TYR;
                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-93295445; PubMed-8515786;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61144 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92.1%;
93.6%;
                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                            LGISFLILMLILYSFI 400
                                                                                                                                                                                                                                                                  LGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity ....
Matches 353; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      539 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PROSITE; PS00109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00008;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro
                                                                                                                                                                                                                                                                                                                                                          025966;
                                                                                                                                                                                                                                                                                                                                              925966
```

```
MEDLINE=93295415; PubMed=8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal part of the precursor to the
major merozoite surface proteins (MSP1) of Plasmodium falciparum from
field isolates.";
GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSNF 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LINTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSNF 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AVTISVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 60
                        ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 5; Length 539;
                                                                                                                                                                                                                                                                                                                                                                                                                                     Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8BF9C64322E9A778 CRC64;
                                                                                                                                                                                                                                                                                                                                                           01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            , 17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 1824.5; Di
Pred. No. 2e-87;
6; Mismatches
                                                                                                                                                                                                                                                                                         Ş
```

1;

61

셤 õ 셤 ŏ

121 284 Ą

3

361

OCC COLOR REPLACE OCC SERVICE OF THE SERVICE OCC SERVI

a ò g

```
GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSNF
                                                                                                                                                                                   MEDLINE-93295445; PubMed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
Sequence conservation in the C-terminal part of the precursor to major merozoite surface proteins (MSP1) of Plasmodium falciparum field isolates.";
                                                                                                                          Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 1824.5; DB 5; Length
Pred. No. 2e-87;
6; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                        C70C2E100EC4A101 CRC64;
                                                  01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
MAJOR WEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
       539
                                                                                                                                                                                                                                                               Mol. Biochem. Parasitol. 59:95-100(1993)
BMBL; D13389; BAA02619.1; -.
InterPro; IPR000561; -.
InterPro; IPR001245; -.
                                                                                                                                                                                                                                                                                                                       Pfam; PF00008; EGF; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; Merozoite; EGF-11ke domain.

1 1 1 SEQUENCE 539 Aa; 61075 MW; C70C2E1
                                      Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                     92.1%;
93.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LGISFLLILMLILYSFI 539
                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 92.1:
Best Local Similarity 93.6'
Matches 353; Conservative
       PRELIMINARY;
                                    (TrEMBLrel.
                                                                                                             Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                      539 AA;
                                                                                                                                                                          SEQUENCE FROM N.A.
                                  01-MAY-2000 (
01-MAY-2000 (
01-JUN-2000 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q25972
Q25972;
      Q9TYG1
Q9TYG1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MSP1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
Q25972
       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Dρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AC DT DT DT OX OX OX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              241 VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 AVTISVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMIFNVNVKDILNSRFNKRE 60
Ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                       MEDLINE-93295445; PubMed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
Sequence conservation in the C-terminal part of the precursor to major merozoite surface proteins (MSPI) of Plasmodium falciparum field isolates.";
                                                                                                                                                                                                                                                              Plasmodium falciparum.
Bukaryota: Alveolata: Apicomplexa: Haemosporida: Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 1824.5; DB 5; Length
Pred. No. 2e-87;
6; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3788015F3127CB9E CRC64;
                                                                                                                                                                        01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
MSP1.
                                                                                                                                                        Ş
                                                                                                                                                                                                                                                                                                                                                                                                                  Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13356; BAA02617.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF00008; EGF; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR;
                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61114 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     92.1%;
93.6%;
                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Merozoite; EGF-like domain.
                                                             523 LGISFLLILMLILYSFI 539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LGISFLLILMLILYSFI 539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 92.1
Best Local Similarity 93.6
Matches 353; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR000561; -. InterPro; IPR001245; -.
                                                 LGISFLLILMLILYSFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AA;
                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                              NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
SEQUENCE
                                                                                                                                                      025984
```

403

Gaps

ä

Length 539;

ţ

462

121

a ò qq ò g ò g ö a ò g à 8

344 181 284

361

RESULT 09TYG1

Ā

us-09-500-376-5.rspt

ä

```
SEOUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROSITE;
                                       NON TER
                                                                                                                                                                                                                                                                                                                                                                                                                                     025971
                                                                                                                                                                                                                                                                                                                                                                                                                                                025971
                                                                                                                                                                                                                                                                                                                                                                                                                  œ
                                                                                                                                                                                 224
                                                                                                                                                                                                                                                                344
                                                                                                                                                             19
                                                                                                                                                                                                     121
                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                     523
                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
 DR DR DR SO
                                                                                                                        à
                                                                                                                                         g
                                                                                                                                                             ò
                                                                                                                                                                               g
                                                                                                                                                                                                     ö
                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                          δy
                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                   엄
                                                                                                                                                                                                                                                                                                                                                                                                                                       1;
                 MEDLINE-93295445; PubMed-8515786; Jongwutiwes S., Tanabe K., Kanbara H.; Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSPI) of Plasmodium falciparum from field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-93295445; PubMed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
Sequence conservation in the C-terminal part of the precursor to the
major merozoite surface proteins (MSP1) of Plasmodium falciparum from
field isolates.";
                                                                                                                                                                                                                                 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                      GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                403
                                                                                                                                                                                                                                                                                                                                                                                     300
                                                                                                                                                                                                                                                                                                                                                                                                                           360
                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                       1 AVTTSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE
                                                                                                                                                                                                                                                                        224 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNJKDSIDTDINFANDVL
                                                                                                                                                                                                                                                                                                                LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                                       GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSNF
                                                                                                                                                                                                                                                                                                                                                                                    VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                     ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                Length
                                                                                                                                                                                                    17; Indels
                                                                                                                                                   539 AA; 61076 MW; 72186F1412C57D1C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-UIN-2000 (TrEMBLrel. 14, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRACHENI)
                                                                                                                                                                                 DB 5;
                                                                                                                                                                                91.9%; Score 1821.5; DB 93.4%; Pred. No. 2.9e-87;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     539 AA
                                                                                                                                                                                                    7; Mismatches
                                                                   Mol. Biochem. Parasitol. 59:95-100(1993)
EMBL; D13360; BAA02621.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mol. Blochem. Parasitol. 59:95-100(1993)
EMBL; D13363; BAA02624.1; -.
                                                                                                     Pfam; PF00008; EGF; PROSITE; PS00109; PROTEIN_KINASE_TYR; Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LGISFLLILMLILYSFI 377
                                                                                                                                                                                          Best_Local Similarity 93.4
Matches 352; Conservative
                                                                                       InterPro; IPR000561; -
InterPro; IPR001245; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR000561;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=5833;
                                                                                                                                                   SEQUENCE
                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                               344
                                                                                                                                                                                                                                                                                                                                                                                                                                             463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    523
                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                         284
                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                   241
                                                                                                                                                                                                                                                                                                                                                                                                                         301
ò
                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                              á
                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                    ç
                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                  ογ
                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
```

```
240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of Plasmodium falciparum from
                                                                                                                                                                                                                                                                                                                           NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV
                                                                                                                                                                                                                                                                                                                                                                                                                             1 AVTISVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                              Length 539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                    17; Indels
                                                                                                    3EED87473EE87B65 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61017 MW; 5CA9C651BB62B5E1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 1818.5; DB 5;
Pred. No. 4.1e-87;
                                                                                                                                                              .,
2
                                                                                                                                                            Score 1821.5; DB Pred. No. 2.9e-87;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        539 AA
                                                                                                                                                                                                    7; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             field isolates.";
Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13359; BAA02620.1; -.
                  Pfam; PF00008; EGF; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PROTEIN_KINASE_TYR; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDILINE-92295445; PubMed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-termina"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                  539 AA; 61046 MW;
                                                                                                                                                            91.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       91.8%;
93.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                            Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                361 LGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 352; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR000561; -. InterPro; IPR001245; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF00008; EGF; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Plasmodium falciparum
IPR001245;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             539 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                   Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PS00109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=5833;
```

```
Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      74134 MW; AA2137E699255150 CRC64;
                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-MAY-2000 (TrEMBLrel. 13, Last annotation update) MEROZOITE SURFACE ANTIGEN 1 (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     84;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  58.9%; Score 1167.5; DB 57.9%; Pred. No. 2.6e-53; ive 59; Mismatches 84
                                                                                                                                                                                                                       651 AA
                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                     MEDLINE=92275047; Pubmed=1592091;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     merozoite surface antigen-1.";
Exp. Parasitol. 81:47-54(1995).
EMBL; Z35329; CAA84558.1; -.
                                                                                                                                                                                                                                          (TrEMBLrel. 01,
                                                                                                                                                               LGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches 226; Conservative
                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                              Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR000561; - Pfam; PF00008; EGF; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       651 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                              [2]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                         01-NOV-1996
01-NOV-1996
                                                                                                                                                                                                                                                                                                                            STRAIN=RO-71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Merozoite.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                     025924
025924;
                                                                                                                                                     361
                                                                                                                                                                                                    RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              67
                                     181
                                                                           241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     127
                                                                                                               301
                                                                                                                                                                       523
                QQ
                                     à
                                                      Db
                                                                         φ
                                                                                            g
                                                                                                               φ
                                                                                                                                 q
                                                                                                                                                    οy
                                                                                                                                                                      g
                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ï
                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
MEDLINE=93295445; PubMed=8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSPl) of Plasmodium falciparum from field isolates.";
                                    223
                                                                 180
                                                                                                               343
                                                                                                                                  240
                                                                                                                                                     403
                                                                                                                                                                      300
                                                                                                                                                                                          462
                                                                                                                                                                                                            360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                  9
 Gaps
                 NFKNVLESDLIPYKDLTSSNY•VVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL
                                                                                                       LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                             VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                 GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSNF
                                                                                                                                                                                                                     1 AVTTSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV
1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ÷
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 539;
                                                                                                                                                                                                                                                                                                                                                                          Plasmodium falciparum.
Eukaryota: Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            105C2E02974FB398 CRC64;
                                                                                                                                                                                                                                                                                                                                           01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-JUN-2000 (TrEMBLrel. 14, Last annotation update) MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 1818.5; DB
Pred. No. 4.1e-87;
6; Mismatches 18
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mol. Blochem. Parasitol. 59:95-100(1993).
EMBL; D13361; BAA02622.1; -
InterPro; IPR000561; -.
InterPro; IPR001245; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROSITE; PS00109; PROTEIN_KINASE_TYR; 1
                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                   PRT;
9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61047 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ouery Match 91.8%;
Best Local Similarity 93.4%;
Matches 352; Conservative
                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                     01-NOV-1996 (TrEMBLrel. 01,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Merozoite; EGF-like domain.
NON_TER 1 1 SEQUENCE 539 AA; 61047 M
                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PF00008;
Matches 352;
                                                                                                                                                                                                                                                                                                                  025973
                                                                                                                                                                                                                                                                                               σ
                                                                         224
                                                                                                               284
                                                                                                                                                                      241
                                                                                                                                                                                        404
                                                                                                                                                                                                           301
                                                                                                                                                                                                                              463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                164
                                                                                            121
                                                                                                                                181
                                                                                                                                                   344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      224
                                                       19
                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               요
                                                                                                                                                 q
                                                                                                                                                                                        g
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ġ.
                   ò
                                   g
                                                      ò
                                                                       g
                                                                                            ô
                                                                                                              g
                                                                                                                                ò
                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
343
                                                                                                                                                                                                                                           462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 of
284 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV
                                                                             LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                       LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                       VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                       463 GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7 IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEKYKSDLDSIKKYI-------NDKQGENEKYLPFLNNIETLYKTVND
                                                                                                                                                                                                                                                                                                 GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-95354793; PubMed-7628566;
Tolle R., Bujard H., Cooper J.A.;
"Plasmodium falciparum: variations within the C-terminal region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 651;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Olafsson P., Matile H., Certa U.;
"Plasmodium falciparum: the repetitive MSA-1 surface protein RO-71 isolate is recognized by mouse antibody against the nonrepetitive repeat block of RO-33.";
Exp. Parasitol. 74:381-389(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Alveelata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
```

9

```
δ
                 g
                                                                       엄
                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δy
                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                           ä
                                   287
                                                                       347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        561
                                                                                        621
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYN 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7 IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL 66
                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
KIDLFVIHLEAKVLNYTYEKSNVĖVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYN
          504 KKDMLGKLLSTGLV-QNPPNTIISKLIEGKF-QDMLNISQHQCVKKQCPQNSGCFRHLDE
                                                                      HNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTM1QQSSGCFRHLDE
                                   HNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDE
                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDLINE=93295445; PubMed=8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
Sequence conservation in the C-terminal part of the precursor to major merozoite surface proteins (MSP1) of Plasmodium falciparum field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                          21;
                                                                                                                                                                                                                                             Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                         Length
                                                                                                                                                                                                                                                                                                                                                                                                                           84; Indels
                                                                                                                                                                                                                                                                                                                                                                               A26C45914AF01396 CRC64;
                                                                                                                                                                                     01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT).
Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                        Score 1164.5; DB 5;
Pred. No. 3.2e-53;
                                                                                                                                                                        ₹
                                                                                                                                                                                                                                                                                                                                                                                                                         60; Mismatches
                                                                                                                                                                        269
                                                                                                                                                                                                                                                                                                                         Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13346; BAA02607.1; -.
Interpro: IPR000561; -.
Pfam; PF00008; EGF; I.
Merozoite; EGF-like domain.
                                                                                                          PLFDGIFCSSSNFLGISFLLILMLILYSFI 377
                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                               64485 MW;
                                                                                                                                                                                                                                                                                                                                                                                                        58.8%;
57.7%;
                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 57.7'
Matches 225; Conservative
                                                                                                                                                                                                                                                                                                                                                                              569 AA;
                                                                                                                                                                                                                                                    NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                       NON_TER
SEQUENCE
                                                                                                                                                                      025969
                444
                                                                                       562
                                                                                                         348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      228
                                  228
                                                                                                                          622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        422
                                                                      288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          480
               g
                                   ò
                                                   a
                                                                      ò
                                                                                      В
                                                                                                          å
                                                                                                                          g
                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
```

```
ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----NDKQGENEKYLPFLNNIETLYKTVND 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYN 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCY 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7 IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL 66
                                                                                                                                                                                                                                                                                                                                             Ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  major merozoite surface proteins (MSP1) of Plasmodium falciparum field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21;
                                                                                                                                                                                                                  Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        84; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64457 MW; A2691FE569E80396 CRC64;
                                                                                                                                            01-NOV-1996 (TrEMBLrel. 01, Created)
1-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 1164.5; DB Pred. No. 3.2e-53;
                                                                                                                 Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ΑA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      60; Mismatches
                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
MEDILINE=93295445; Pubmed=8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              269
                                                                                                                                                                                                                                                                                                                                                                                         Mol. Biochem. Parasitol. 59:95-100(1993)
EMBL; D13348; BAA02609.1; -
377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PLFDGIFCSSSNFLGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (TrEMBLrel. 01, Created)
(TrEMBLrel. 01, Last sequ
(TrEMBLrel. 13, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PLFDGIFCSSSNFLGISFLLILMLILYSFI
               PLFDGIFCSSSNFLGISFLLILMLILYSFI
PLFDGIFCSSSNFLGISFLLILMLILYSFI
                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEKYKSDLDSIKKYI -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    58.8%;
                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 225; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR000561;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         025975
025975;
01-NOV-1996 (
01-NOV-1996 (
01-MAY-2000 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local (
                                                                                                               025974
348
                                540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
Q25975
                                                                               RESULT
Q25974
```

ä,

Gaps

21;

361

```
SEQUENCE FROM N.A.
MEDLINE-93295445; PubMed-8515786;
Jongwatiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates.";
                                                                                                                                                                                                                                                                                                       ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEKYKSDLDSIKKYI--------NDKOGENEKYLPFLNNIETLYKTVND 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                168 KIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYN 227
                                                                                                                                                                                                                                                                                                                                                                                                7 IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    228 HNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         288 REECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 1164.5; DB 5; Length 569;
Pred. No. 3.2e-53;
); Mismatches 84; Indels 21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                               Length
                                                                                                                                                           58.8%; Score 1164.5; DB 5; Length 57.7%; Pred. No. 3.2e-53; Live 60; Mismatches 84; Indels
                                                                                                            A26C45914AF01396 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          269
   Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13350; BAA02611.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13352; BAA02613.1; -.
InterPro; IPR000561; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             348 PLFDGIFCSSSNFLGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PLEDGIFCSSSNFLGISFLLILMLILYSFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .09
                                                                                                        64485 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64457 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  58.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                    InterPro; IPR000561; -. Pfam; PF00008; EGF; 1. Merozoite; EGF-like domain.
                                                                                                                                                                               Best Local Similarity 57.7
Matches 225; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 58.8
Best Local Similarity 57.7
Matches 225; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Plasmodium falciparum
                                                                                                            569 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ¥
                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NON_TER
SEQUENCE
                                                                                                            SEQUENCE
                                                                                                                                                             Query Match
                                                                                           NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      025979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                540 ]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 15
                                                                                                                                                                                                                                                                                                       67
                                                                                                                                                                                                                                                                                                                                                                            127
       g
                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                               3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-93295445; PubMed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal part of the precursor to the
major merozoite surface proteins (MSP1) of Plasmodium falciparum from
field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                127 SEKYKSDLDSIKKYI-------1000GENEKYLPFLNNIETLYKTVND 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KIDLEVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYN 227
                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                              7 IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL 66
                                                                                                                                                               ಭ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 228 HNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDE
                                                                                                                                         Jongwutiwes S., Tanabe K., Kanbara H.;
Sequence conservation in the C-terminal part of the precursor to
major merozoite surface proteins (MSP1) of Plasmodium falciparum
field isolates."
                                                                                                                                                                                                                                                                                                                                                                                                               21;
                                                                                                                                                                                                                                                                                                                                                                         58.8%; Score 1164.5; DB 5; Length 569; 57.7%; Pred. No. 3.2e-53; 1ve 60; Mismatches 84; Indels 21;
                                  Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                                     ED4284B2867C9703 CRC64;
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1996 (TrEMBLrel. 01, Created)
U-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ş
                                                                                                                                                                                                               Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13349; BAA02610.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             348 PLFDGIFCSSSNFLGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                          MEDLINE=93295445; PubMed=8515786;
                                                                                                                                                                                                                                                                                                                     569 AA; 64459 MW;
                                                                                                                                                                                                                                                                Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 57.7
Matches 225; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MSP1.
Plasmodium falciparum.
                                                                                                                                                                                                                                                InterPro; IPR000561;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                        SEQUENCE FROM N.A.
                                                                     NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                     NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                          Query Match
```

242

29

ò g ô g ò 셤 ò g ŏ 8 ò g

168

025977

RESULT 14

3;

Gaps

```
ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL 126
                                                                                                    ----NDKQGENEKYLPFLNNIETLYKTVND 167
                                                          KIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYN 227
7 IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL 66
                                                                                                                                                      127 SEKYKSDLDSIKKYI---
                                     242
                                                                           168
                                                                                                                                                       348
                                                                                                                                                                   540
                         67
                                                              302
                                                                                                                g
           g
                         ŏ
                                    Db
                                                 ò
                                                              g
                                                                           ò
                                                                                      g
                                                                                                    å
                                                                                                                             õ
                                                                                                                                         q
                                                                                                                                                       ò
```

Search completed: August 8, 2001, 12:33:47 Job time: 275 sec

```
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model Run on:

August 8, 2001, 12:32:38 ; Search time 24.79 Seconds (without alignments) 520.949 Million cell updates/sec

US-09-500-376-5 1982 1 AVTTSVIDNILSKIENEYEV......SNFLGISFLLILMLILYSFI 377 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

93435 segs, 34255486 residues Searched:

93435 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_39:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description							P08569 plasmodium					P33539 agaricus bi		Q58285 methanococc		P75093 mycoplasma		-			Q49419 mycoplasma	_	O51578 borrelia bu	O05949 rickettsia	Q09622 caenorhabdi		,	-	P27895 saccharomyc			P13823 plasmodium	P40957 saccharomyc
SUMMARIES	Ω	MSP1_PLAFK	MSP1_PLAFW	MSP1_PLAFC	MSP1_PLAFP	MSP1_PLAF3	MSP1_PLAFF	MSP1_PLAFM	MSP1_PLAYO	RBP2_PLAVB	Y704_METJA	YLJ2_CAEEL	RPOP_AGABT	IRR1_YEAST	Y875_METJA	CHEA_BORBU	Y018_MYCPN	LHS1_YEAST	BXEN_CLOBU	BXCN_CLOBO	ATX1_PLAFA	X328_MYCGE	BXEN_CLOBO	EX5B_BORBU	DPO1_RICPR	YS83_CAEEL	NUF1_YEAST	HS83_PHANI	YKDA_MYCCA	CIN8_YEAST	YCF1_TOBAC	RBP1_PLAVB	SERA_PLAFG	MAD1_YEAST
	DB	ู่ ส	Н	٦	Н	Н	П	-	П	-	-	-	-	П	Н	1	-	Н	-	-	-	-	_	-4	-	-	7	٦	7	-	-	П	-	П
	Query Match Length	1630	1639	1726	1726	1682	1701	1701	1772	1251	377	2198	1102	1150	748	864	1030	881	1162	1196	1956	756	1162	1169	867	166	944	703	655	1038	1901	2869	989	749
d₽	Query	92.1	92.1		58.7	•			٠.	•		6.9	•	9.8	6.7	6.7	٠	٠	٠	٦.	•	•		6.5				•		٠	٠	6.3	6.2	6.2
	Score	1824.5	1824.5	1164	1163	1156.5	1145	1145	588.5	142	139.5	136.5	135	134	133	133	133	130.5	130	130	129.5	129	129	129	128.5	127	.126.5	126	125.5	125	125	124.5	2	122.5
	Result No.	Т	7	m	4	ស	9	7	8	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	22	56	27	28	29	30	31	32	33

Q08372 plasmodium P75440 mycoplasma	P39682 saccharomyc O51246 borrelia bu	P19214 plasmodium P25386 saccharomyc	P39526 saccharomyc O14525 homo sapien	010310 schizosacch	Q60563 mesocricetu P05475 kluveromvo	o Imparo i Francisco
S230_PLAFO Y242_MYCPN	PR39_YEAST Y228_BORBU	EBA1_PLAFC USO1_YEAST	YJU7_YEAST	YD56_SCHPO	SCP1_MESAU	
						1
3135 632	629 971	1435	2014	515	845	
6.2	6.1	0.9	6.0 6.0	0.9	0.9	
122	120.5 119.5	119.5	119.5 119	118.5	118.5	)
34 35	36 37	388	40	4 4 2 8	4 4 5	

## ALIGNMENTS

us-09-500-376-5.rsp

```
1;
                                                                                                                                                                                                                                                                                                                     1495 VFENLAKTVLSNLLDGNL-QGMLNISQHQCVKKQCPQNSGCFRHLDEREECKCLLNYKQE 1553
                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-86014355; PubMed-2995820;
Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V.,
Nicholls S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T.,
                                                                                                                                                                                                                     NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                     GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                       VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE 300
                                                                                                                                                                                                                                                                                                                                                                                                                         GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSNF 360
                                                                                                                                            Gaps
                                                                                                                                                                 1 AVTISVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Holder A.A.;
Submitted (MAR-1991) to the EMBL/GenBank/DDBJ databases.
-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PTM: MEROZOTTE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ol-FEB-1996 (Rel. 33, Last sequence update)
01-FEB-1996 (Rel. 33, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
(PRMSA) (P195).
                       (POTENTIAL)...) (POTENTIAL)...) (POTENTIAL)...) (POTENTIAL)...)
              (POTENTIAL)
                                                                                                                92.1%; Score 1824.5; DB 1; Length 1630; 93.6%; Pred. No. 2.5e-90;
                                                                                                                                           ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                antigens of Plasmodium falciparum merozoites.";
Nature 317:270-273(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Plasmodium falciparum (isolate Wellcome).
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                         17; Indels
                                                                           ADBDEC3CE0A46322 CRC64;
                         (GLCNAC. . .)
N-LINKED (GLCNAC. N-LINKED (GLCNAC. N-LINKED (GLCNAC. N-LINKED (GLCNAC. N-LINKED (GLCNAC. N-LINKED (GLCNAC. N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT; 1639 AA
                                                                                                                                         6; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-AUG-1987 (Rel. 05, Created)
                                                                           187289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1614 LGISFLLILMLILYSFI 1630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         361 LGISFLLILMLILYSFI 377
                                                                                                                                        353; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
955
1049
1156
1165
1436
                                                                            AA;
                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
955
1049
1156
1165
1165
1436
1517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=5848;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MSP1_PLAFW
P04933;
                                                                                                                Query Match
Best Local Si
Matches 353;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REVISIONS
                                   CARBOHYD
CARBOHYD
           CARBOHYD
                         CARBOHYD
CARBOHYD
                                                             CARBOHYD
                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PLAFW
                                                                                                                                                                                           1255
                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                               1435
                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                        301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FT FT SS
                                                                                                                                                                                           엄
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                임
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                         ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                             NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                         1; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                1 AVTTSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSNF
                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                 Length 1639;
                                                                                                                              InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                        17; Indels
                                                                                                                                                                                                                                                                                                                                                               CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                  1;
                                                                                                                                                                                                                                                                                                                                                   LINKED (GLCNAC. . . . 2C255B6616C87F6E
                                                                                                                                                                                                                                                                                                                                                                                                 Score 1824.5; DB Pred. No. 2.6e-90;
                                                                                                                                                                             POTENTIAL.

MEROZOITE SURFACE
N-LINKED GLCNAC.
                                                                                                                                                                                            SURFACE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                         6; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 1726 AA.
                                                                                                                                                                                                                                                                     20 920
64 964
8 1058
1165
1174
1445
1 1826
1 187618 M
                                                                                                                                                                                                                                                                                                                                                                                                 92.1%;
93.6%;
                                                                                                         EMBL; X02919; CAA26676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LGISFLLILMLILYSFI 1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LGISFLLILMLILYSFI 377
                                                                                                                                                                    Transmembrane; GPI-anchor
                                                                                                                                                                                                                                                                                                                                                                                                                         Matches 353; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               05,
10,
34,
                                                                                                                      PIR; A24594; A24594.
                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (Rel. (Rel. (Rel. (Rel. )
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-MAR-1989
01-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               13-AUG-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MSP1_PLAFC
P04934;
                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local (
                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                 SIGNAL
                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 3
MSP1_PLAFC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
```

```
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               isolate.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIGNAL
                                                   1518
                                                                                                                                                                                                                                                                                                                             MSP1_PLAFP
                   167
                                                                                     227
                                                                                                                                                                                                                         347
                                                                                                                                                         287
                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                  a
                                                                                                                                                                                       g
                                                                                                                                                                                                                                                        qq
                                                                                     ò
                                                                                                                                                       ŏ
                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1104-1726 FROM N.A.
MEDLINE-88143999; PubMed=3278296;
Weber J.L., Sim B.K.L., Lyon J.A., Wolff R.;
Werozoite surface protein sequence from the Camp strain of the human malaria parasite Plasmodium falciparum.";
Nucleic Acids Res. 16:1206-1206(1988).
-:- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-86205236; PubMed-3517809;
MEDLINE-86205236; PubMed-3517809;
Webber J.L., Leininger W.M., Lyon J.A.;
Warlation in the encoding a major merozoite surface antigen of the human malaria parasite plasmodium falciparum.";
Nucleic Acids Res. 14:3311-3323(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7 IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL 66
                                                                                                                                                                                                                                                                                                                                                                                                 Q.
 MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                 ANTIGENS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22;
                                                                                                                                                                                                                                                                                                                                                                               83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        58.7%; Score 1164; DB 1; Length 1726; 57.5%; Pred. No. 4.7e-55; 1ve 60; Mismatches 84; Indels 22;
                                                  p / Malaysia).
Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                          PTM: MEROZÓTTE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANT MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SURFACE PROTEIN 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC...)
GLCNAC...)
GLCNAC...)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC. ..)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DD8AD45FA352BCF3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL.
MEROZOITE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-L
                                                  (isolate Camp
Apicomplexa; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; X03831; CAA27446.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       196197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GPI-anchor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                  Plasmodium falciparum
Eukaryota; Alveolata;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SAZOGM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 Similarity
225; Conserv
                                                                                   NCBI_TaxID=5835;
                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                PMMSA) (P195).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR; A23386;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1338 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               67
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                              EREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDC 346
                                                                                                                                             DKIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDY 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "Plassmodium falciparum: gene structure and hydropathy profile of the major merozoite surface antigen (gpl95) of the Uganda-Palo Alto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-0CT-1996 (Rel. 34, Last sequence update)
01-0CT-1996 (Rel. 34, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EXP. Parasitol. 67:1-11(1988).
-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                           NHNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL).
PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OMEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Plasmodium falciparum (isolate Palo Alto / Uganda).
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-89005525; Pubmed-3049134;
Chang S.P., Kramer K.J., Yamaga K.M., Kato A., Case S.E.,
Siddiqui W.A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PROTEIN 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL.

MEROZOITE SURFACE
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SURFACE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 1726 AA
                                                                                                                                                                                                                                                                                                                                                      1696 YPLFDGIFCSSSNFLGISFLLILMLILYSFI 1726
                                                                                                                                                                                                                                                                                                                            YPLFDGIFCSSSNFLGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; M37213; AAA29611.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Transmembrane; GPI-anchor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1726
133
272
272
501
567
638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=57270;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (PMMSA) (GP195).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-0CT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MSP1_PLAFP
P50495;
```

```
Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MSP1_PLAFF
P13819;
                                                                                                                                                                                                                                                                                                                                                            Best Local Sim
Matches 225;
  use by n
modified a
entities r
                                                                                                                                                                CHAIN
TRANSMEM
                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (PMMSA)
                                                                                                                                                   SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MSP1_PLAFF

MSP1_PLAFF

1D MSP1_P

AC P13819

DT 01-JAN

DT 01-JAN

DD MENCOZO

DE MERCOZO

DE MERCOZO

DE MSP-1.
                                                                                                                                                                                                                                                                                                                                                                                                                  1295
                                                                                                                                                                                                                                                                                                                                                                                                                                                              1355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1653
                                                                                                                                                                                                                                                                                                                                                                                                                                         67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     127
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                     3;
                                                                                                                                                                                                                                                                                        1636 EREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDS 1695
                                                                                                                                                    SEKYKSDLDSIKKYI---------NDKQGENEKYLPFLNNIETLYKTVN 166
                                                                                                                                                                                                                                                                            NHNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLD 286
                                                                                                                                                                                                                                                                                                                       EREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDC 346
                                                                                                                                        ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL 126
                                                                      Gaps
                                                                                            99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-8816657; Pubmed-3327688;
Certa U., Rotmann D., Matile H., Reber-Liske R.;
"A naturally occurring gene encoding the major surface antigen precursor pl90 of Plasmodium falciparum lacks tripeptide repeats.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-FEB-1991 (Rel. 17, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
(PMMSA) (P190)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  merozoite surface antigen-1.";
Exp. Parasitol. 81:47-54(1995).
-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ot
                                                                                           7 IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL
                                                                                                                                                                                                                                DKIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PTM: MEROZÓTTE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1032-1682 FROM N.A.
MEDILIRE-95354793; PubMed-7628566;
Tolle R., Bujard H., Cooper J.A.;
"Plasmodium falciparum: variations within the C-terminal region
 1613 N-LINKED (GLCNAC. . .) (POTENTIAL)
AA; 196174 MW; 5B59CEEFA2F9A026 CRC64;
                                                                    22:
                                               Length 1726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
NCBI_TaxID=5834;
                                                                     84; Indels
                                             Score 1163; DB 1;
Pred. No. 5.3e-55;
60; Mismatches 84;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    1682 AA
                                                                                                                                                                                                                                                                                                                                                                                 1696 YPLFDGIFCSSSNFLGISFLLILMLILYSFI 1726
                                                                                                                                                                                                                                                                                                                                                                     347 YPLFDGIFCSSSNFLGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                            58.78;
57.58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     precursor p190 of Plasmod1 EMBO J. 6:4137-4142(1987).
                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                  MSP1_PLAF3
P19598; Q25921;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                    225;
                                             Query Match
Best Local S:
Matches 225
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 5
MSP1_PLAF3
                                                                                                                                                            1398
                                                                                                                                                                                                          1458
                                                                                                                                                                                                                                                     1518
                                                                                                                                                                                   127
                                                                                                                                                                                                                                167
                                                                                                                                                                                                                                                                            227
                                                                                                                                       67
 E S
                                                                                                                g
                                                                                                                                       ò
                                                                                                                                                            g
                                                                                                                                                                                    õ
                                                                                                                                                                                                         a
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                          셤
```

```
modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3,
    in no way
commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNPYDFEAIKKLINDDT 1534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEKYKSDLDSIKKYI---------NDKQGENEKYLPFLNNIETLYKTVND 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYN 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7 IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-JRN-1990 (Rel. 13, Created)
1-JRN-1990 (Rel. 13, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REECKCLLNYKOEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKK1TCECTKPDCY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   58.4%; Score 1156.5; DB 1; Length 1682; 57.7%; Pred. No. 1.1e-54; ive 58; Mismatches 86; Indels 21;
                                                                                                                                                                                                                                                                                                                                                               . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                    Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
  as its content
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C82A1E159948CAD6
                                                                                                                                                                                                                                                                                          POTENTIAL.
MEROZOTTE SURFACE PI
MEMBRANE ANCHOR (PO'
N-LINKED (GLCNAC.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 1701 AA
non-profit institutions as long
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PLFDGIFCSSSNFLGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              58;
                                                                                                      EMBL; M35727; AAA29715.1; --
EMBL; Y00087; CAA68280.1; --
EMBL; 235326; CAA84555.1; --
PIR; S06286; S06286.
InterPro; IPR000561; --
Pfam; PF00008; EGF; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   192462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
```

```
3,
                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       between the Swiss Institute of Bioinformatics and the EMBL outstation—the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7 IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL 66
                                                                                                                                                     Peterson M.G., Coppel R.L., McIntyre P., Langford C.J., Woodrow G., Brown G.V., Anders R.F., Kemp D.J.;
Wariation in the precursor to the major merozoite surface antigens of Plasmodium falciparum:
                                                                                                                                                                                                                                                                                                                                       -!- PTM: MEROZÓITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 4 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                           Mol. Biochem. Parasitol. 27:291-302(1988).
-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEKYKSDLDSIKKYI------NDKQGENEKYLPFLNNIETLYKTVN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DKIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
Malaria, Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
Transsmembrane; GPI-anchor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           57.8%; Score 1145; DB 1; Length 1701; 56.8%; Pred. No. 4.7e-54;
Plasmodium falciparum (isolate FC27 / Papua New Guinea).
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SURFACE PROTEIN 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3920B75E73D38552 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC...)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              86;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEROZOITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                  MEDLINE-88142999; PubMed-2449612;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M19143; AAA29653.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               193719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              222; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1089
1196
1588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  239
470
536
607
802
899
919
991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PIR; A54498; A54498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                      SEQUENCE FROM N.A.
                                                NCBI_TaxID=5837;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local
Matches 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
SOURCE STATE STATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

NHNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLD

227

ò

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseelisb-sib.ch).
                              MEDLINE-88011243; PubMed-3079521;
Tanabe K., Mackay M., Goman M., Scaife J.G.;
"Allelic dinorphism in a surface antigen gene of the malaria parasite
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        42
EREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDC 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q.
                                                                                                                                                                                                                                                                                                                                               30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
(PPMSA) (P190)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U., Stunnenberg H., Bujard H.;
Stunnenberg H., Bujard H.;
Folymorphism of the precursor for the major surface antigens of Plasmodium falciparum merozoites: studies at the genetic level.";
EMBO J. 4:3823-3829(1985).
-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        83 KDA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURPACE ANTIGENS MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
(POTENTIAL)
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Guinea).
Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROTEIN 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Tanabe K.;
Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (isolate mad20 / Papua New Apicomplexa; Haemosporida;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL.
MEROZOCITE SURFACE PI
N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                1701 AA
                                                                                                                                      1671 YPLFDGIFCSSSNFLGISFLLILMLILYSFI 1701
                                                                                                          347 YPLFDGIFCSSSNFLGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REVISIONS TO 1403; 1569 AND 1629.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 1-115 FROM N.A. MEDLINE-86136024; PubMed=3004972;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Plasmodium falciparum.";
J. Mol. Biol. 195:273-287(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; X05624; CAA29112.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GPI-anchor
                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Alveolata;
NCBL_TaxID=70153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A26868; A26868
B25120; B25120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1110
2339
470
536
607
802
899
919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Transmembrane;
                                                                                                                                                                                                                                                                                                MSP1_PLAFM
P08569;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIGNAL
                                                                                                                                                                                                                                             RESULT 7
MSP1_PLAFM
287
ð
                                               g
                                                                                                       ò
                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                   HID DEBAR SERVICE SECOND SERVICE SERVI
```

9

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                 1610
                                                                                                                                                                                                                                                                                                                                                                                                                                                      1670
                                                                                                                                                                                                                        STRAIN=17XL;
MEDLINE=8812489; PubMed=2448778;
BURDS J.M. Jr., Daly T.M., Vaidya A.B., Long C.A.;
Burns J.M. Jr., Daly T.M., Vaidya A.B., Long C.A.;
"The 3' portion of the gene for a Plasmodium yoelii merozoite surface antigen encodes the epitope recognized by a protective monoclonal antibody.";
                                                                                                                                                                                                                                                                                                                                                                                                                            346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      the
                                                                                                                         Gaps
                                                                                                                                                    99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL).
PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 4
KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    90 (Rel. 13, Created)
90 (Rel. 13, Last sequence update)
96 (Rel. 34, Last anotation update)
SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Proc. Natl. Acad. Sci. U.S.A. 85:602-606(1988).
                                                                                                                                                                                                       ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL
                                                                                                                                                                                                                                                                           1553 TKKDMLGKLLSTGLV-QNPPNTIISKLIEGKF-QDMLNISQHQCVKKQCPENSGCFRHLD
                                                                                                                                                                                                                                                                                                                                                                                                                                      ţ
                                                                                                                                                 7 IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL
                                                                                                                                                                                                                                                             ------NDKQGENEKYLPFLNNIETLYKTVN
                                                                                                                                                                                                                                                                                                                  DKIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDY
                                                                                                                                                                                                                                                                                                                                                                      NHNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLD
                                                                                                                                                                                                                                                                                                                                                                                                                           EREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDC
N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
MW; 3FC2EC59AF96EA98 CRC64;
                                                                                                                       22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Plasmodium yoelii.";
                                                                                             Length 1701;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gene encoding the
                                                                                            Score 1145; DB 1; DPred. No. 4.7e-54; 61; Mismatches 86;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   major merozoite surface antigens of Plasmo
Mol. Biochem. Parasitol. 39:285-288(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1671 YPLFDGIFCSSSNFLGISFLLIALLYSFI 1701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                YPLFDGIFCSSSNFLGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     'Sequence analysis upstream of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=90205979; PubMed=2320061;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1093-1772 FROM N.A.
                                                                                            57.8%;
56.8%;
                                                    193768
                                                                                                                                                                                                                                                             SEKYKSDLDSIKKYI-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Plasmodium berghei yoelii.
                                                                                           Query Match 57.8°
Best Local Similarity 56.8
Matches 222; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
991
1089
1196
1588
                                                       ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (PMMSA) (230 KDA)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=5862;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-JAN-1990 (
01-JAN-1990 (
01-OCT-1996 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MSP1_PLAYO
P13828;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Lewis A.P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEROZOITE
                                      CARBOHYD
SEQUENCE
            CARBOHYD
CARBOHYD
                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 8
MSP1_PLAYO
                                                                                                                                                                            1313
                                                                                                                                                                                                                                  1373
                                                                                                                                                                                                                                                                                     1433
                                                                                                                                                                                                                                                                                                                                          1493
                                                                                                                                                                                                                                                                                                                                                                     227
                                                                                                                                                                                                                                                                                                                                                                                                                                                     1611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              347
                                                                                                                                                                                                       29
                                                                                                                                                                                                                                                           127
                                                                                                                                                                                                                                                                                                                167
                                                                                                                                                                                                                                                                                                                                                                                                                           287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
FT
                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                   유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDD
                                                                                                                                                                            셤
                                                                                                                                                                                                       ò
                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                          á
                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DNYKNEKAEFEILTKNLEKYIQIDEKLDEFVEHAENNKHIASIA------LUNLN 1629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1630 KSGLVGEGESKKILAKMINMDGMDLLGVDPKHVCVDTRD---IPKNAGCFRDDNGTEEWR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    :| :||:::||:::||:::||:::|| : : : | |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |: :| |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9 NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVLES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KYKSDLDSIKKYI-----ND-KQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1521 LYNTQLAAVKEQIATIEAETNDTNKEEKKKYIPILEDLKGLYETVIGQAEEYSEELQNRL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LNYTYEKSNVEVKIKELNYLKTIQDKLADF - - - - KKNNNFVGIADLSTDYNHNNLTKFL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CLLNYKQ - EGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STGMVFENLLKSVLSNLLDWK----LARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1772;
                                                                                                                                                                                                                                                                                                                                   Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          V (IN REF. 2).
9A6291658EB0F45D CRC64:
                                                                                                                                                                                                                                                                                                                                                                                            MEROZOTTE SURFACE PI
N-LINKED (GLCNAC.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 79; Mismatches 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2.4e-24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 588.5;
Pred. No. 2.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1251
                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GIFCSSSNFLGISFLLILMLILYS 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Created)
                                                                                                                                                                                                                                   EMBL, J03612; AAA29762.1; -. EMBL, J04668; AAA29702.1; -. PIR; A28121; A28121. PIR; A45532; A45532.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29.78;
34.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA; 197230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches 132; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        646
829
1018
1090
1408
1446
1541
1680
1521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (Rel.
(Rel.
(Rel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1541
1629
1680
1521
1772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1090
                                                                                                                                                                                                                                                                                                                                                         Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-APR-1993
01-APR-1993
01-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PLAVB
                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RBP2_PL
Q00799;
                                                                                                                                                                                                                                                                                                                                                                               SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 9
RBP2_PLAVB
ID RBP2_PI
AC Q00799;
DT 01-APR-
DT 01-OCT-
                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   237
ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       òγ
```

```
NCBI_TaxID=6239;
                                                                                                                                                                                                                                                                                                                                                                                                                                           57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              YLJ2_CAEEL
P34367;
                                                                                                                                                         jannaschii
                                                                                                                                                                                                                                                                                                                                                            TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                  TRANSMEM
                                                                                                                                                                                                                                                                                                               TIGR;
                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 126
  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                              15;
                                                                                                                                        -!- FUNCTION: INVÒLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO HUMAN RETICULOCYTE CELLS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1053 VTDSKTIISKLKGVIIEVNENTEMNTIESSAKEIEALYNELKNKKTSLNEIYQTSN-EVK 1111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           993 LINTIEKNVKAYLAYIKKNYEDTVQDVLTLNEHFNTKQVSNHEPTNFDKSNKSSEELTKA 1052
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   59 RENFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSI-----DTD 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----LNNIETLYKTVNDKIDLFVIHLEAK-------VLNYTYEKSNVEVK 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            194 IKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGMYFENLLKSVLSNL 253
                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                    3 TTSVIDNILSKIENEYEVL--YLKPLAGVYRSLKKOLEN--NVMTFNVNVKDILNSRFNK:58
                                                                                                                                                                                                                                                                                                                                                                                                                                  88;
                  Plasmodium vivax (strain Belem).
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=31273;
                                                                                                                                                                                                                                                                                                                                                                       Length 1251;
                                                                                           Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.; "A reticulocyte-binding protein complex of Plasmodium vivax
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MJ0704.
Methanococcus jannaschii.
Archaea; Euryarchaeota; Methanococcales; Methanococcaee;
Methanococcus.
                                                                                                                                                                                                                                                                                                                                                                      7.2%; Score 142; DB 1; Length 12
23.2%; Pred. No. 1.1;
tive 53; Mismatches 117; Indels
                                                                                                                                                                                                                                                                                                                                   1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                113 I -- NFANDVLGYYKILSEKYK -- - SDLDSIKKYINDKQGENEKYLPF ---
                                                                                                                                                                  -! - SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1162 IDADSSFTLESIKKFNEIYSHIKTNIGELEQLQQTN 1197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          254 LD------WKLARYVKHFTTPMRKKTMIQQSS 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (Rel. 35, Created)
(Rel. 35, Last sequence update)
(Rel. 35, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ą
RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).
                                                                                MEDLINE=92315338; PubMed=1617731;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HYPOTHETICAL PROTEIN MJ0704.
                                                                                                                                                                                                                                                                                                   Malaria; Receptor; Membrane
                                                                                                                                                                                                                                                                                     EMBL; M88098; AAA29744.1;
                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 23.28
Matches 78; Conservative
                                                                                                                               69:1213-1226(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=2190;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Y704_METJA
Q58115;
                                                                                                                                                                                                                                                                                                             NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          950
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | METJA
엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OC OC OC OC OC
                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12;
SEQUENCE FROM N.A.

STRAIN-JAL-1 / DSM 2661 / ATCC 43067;

MEDLINE-96337999; PubMed-8688087;

Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
"Complete genome sequence of the methanogenic archaeon, Methanococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      86 PYKFLNKEKRDKFLSSYNYIKDSIDTD----INFAND-------VLGYYKI 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LSEKYKS-DLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDK--IDLFVIHLEAKVLN 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26 LAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVLESDLIPYKDLTSSNYVVKD 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nematoda, Chromadorea, Rhabditida, Rhabditoidea, inae, Caenorhabditis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Science 273:1058-1073(1996)
-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              7.0%; Score 139.5; DB 1; Length 377; 25.0%; Pred. No. 0.39;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   334 EKYEKEKNENKIDEEIFNELKDELEKLIN-KCNNK---LRELESONNN 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-FEB-1994 (Rel. 28, Created)
01-FEB-1994 (Rel. 28, Last sequence update)
01-RNOV-1997 (Rel. 35, Last annotation update)
HYPOTHETICAL 256.3 KDA PROTEIN C50C3.2 IN CHROMOSOME III.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2746A0A96425C285 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 183 YTYEKSNVEVKIKE--LNYLKTIQDKLADFKKNNNFVGIADLSTDYNH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              63;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hypothetical protein; Transmembrane TRANSMEM 26 46 POTENTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       [1]
SEQUENCE FROM N.A.
STRAIN-BRISTOL N2;
MEDLINE-94150718; PubMed-7906398;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       39;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                44630 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; U67517; AAB98705.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Eukaryota; Metazoa; Nemato
Rhabditidae; Peloderinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Caenorhabditis elegans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  67
108
135
377 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
```

œ

```
Mitochondrion.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ŋ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   287
                                                                                                                                                                                                                                                               ÷
            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a alicense agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15;
Baynes C., Berks M.,
     Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
Bonfield J., Burton J., Conneil M., Copsey T., Cooper J., Coulson A.,
Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
Latrraille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
Parsons J., Percy C., Riffen L., Roopra A., Saunders D., Shownken R.
Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
Walerson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
Wohldman P.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ENFKNYLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDV 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   243 VGMENDN-----LRNIPDPPKNHVIDVEQRFESFQAFVQKWKTDIENSADADKLMKEAENI 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LGYYKILS--EKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLE 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DNILSKIENEYEVLYLKPLAGVYRSLKKQLE------NNVMTFNVNVKDILNSRFNKR 59
                                                                                                                                                        ပ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       155 SNYEKVHETWSNFQEYIDSLHK----KLSENERFQKFVDNAEDLIKWMDDK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TGMVFENLLKSVLSNLLDWK-----LARYVKHFTTPMRK-KTMIQQSSGCFRHLDEREEC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EKEICEKYSKMDFEVMMK---YRKTVEIEMKSVEOR----IKDLKEOF-----
                                                                                                                                                      '2.2 Mb of contiguous nucleotide sequence from chromosome III of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 2198;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                 CD1B2C1092C5EDC8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-FEB-1994 (Rel. 28, Created)
01-FEB-1994 (Rel. 28, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
PROBABLE DNA-DIRECTED RNA POLYMERASE (EC 2.7.7.6).
Agaricus bitorquis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 1;
                                                                                                                                                                                       -1- SIMILARITY: BELONGS TO THE SPECTRIN FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1102 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 66; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6.9%; Score 136.5; 21.3%; Pred. No. 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EIFDEIRKCSMV-WNEISATFQQMLDP------
                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIĀL.
                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00435; spectrin; 12.
PROSITE; PS00018; EF_HAND; I.
Hypothetical protein; Calcium-binding.
CA_BIND 2025 2036 POTENTIAL
SEQUENCE 2198 AA: 256260 MM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       292 KCLLNYKQEGDKCV-ENPNPTCNE 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                298 CCWSSEKIEDLKIMATSDTPDCDE 321
                                                                                                                                                                                                                                                                                                                                         C50C3.2; CE01861.
                                                                                                                                                                                                                                                                                                                          EMBL; L14433; AAA27973.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002017; -. InterPro; IPR002048; -.
                                                                                                                                                                               Nature 368:32-38(1994)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 Similarity
69; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RPOP_AGABT
P33539;
                                                                                                                                                                                                                                                                                                                                          WormPep;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local
Matches 6
                                                                                                                                                                   elegans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEPTE
         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   THE TRANSCRIPTION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY: BELONGS TO THE PHAGE AND MITOCHONDRIAL RNA POLYMERASES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          71 ELNDSVFLIIKDIEKECKYYVSNVLYRSVGSRKNRGREVEWSSYKYNKEFNKVLDKGIIS 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    115 FANDVLGYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVI 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :| :: || | : :|| | : :|| | : :|| | : :|| :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64 NVLESDLIPYKDLTS--SNYVVKDPYKFLNKEK---RDKFLSSYNYIKD---SIDTD-IN 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE-NFK 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         27 SLVDNLFYNLHD-----YAS---QSKHNIIIDDGKVEKLLNLICNVLDRNVN 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genet. 19:495-502(1991).
FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPT
OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      230 NLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDERE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           175 HLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNF-----VGIADLSTDYNHN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     131 INNEVLKFISKEREGYIERVESIAVTVKN-----KILELNNNIAEVLLSIKNKV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     . . . . . . . . . . . . .
                                                                                                                                                                                                                                                                                                                                                                                                                       Robison M.M., Royer J.C., Horgen P.A.;
"Homology between mitochondrial DNA of Agaricus bisporus and ar
internal portion of a linear mitochondrial plasmid of Agaricus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1; Length 1102;
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Agaricales;
Agaricaceae; Agaricus.
NCBL_TaxID=5343;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mismatches 131; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     F72FDD308D1AB5B4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Interpro; IPR002092; -.
Pfam; PF00940; RNA_pol; 1.
PROSITE; PS00489; RNA_POL_PHAGE_2; 1.
Transferase; Transcription; DNA-directed RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6.8%; Score 135; DB 22.5%; Pred. No. 2.2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ECKCLLNYKQEGDKCVENPNPTCNE 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
STRAIN-ATCC 24666 / AG4;
MEDLINE-91347410; PubMed-1879001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BY
BY
BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; X63075; CAA44799.1; -. PIR; S28104; S28104.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     126664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mitochondrion; Plasmid
ACT_SITE 734 734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1102 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUBSTRATES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RNA(N)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Si
Matches 73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
ACT_SITE
SEQUENCE
```

```
Methanococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hypothetical
SEQUENCE 74
                                                                                                                                                                                                                      Y875_METJA
Q58285;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             annaschii
                 ż
                                                                   275
                                                                                                                                                                                                           Y875_METJA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91
               227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                        9
                                                                   Qγ
                                                                                            g
                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         949 PADATTWRYSMSEQFPFYLHDNASKVLLKIFLYLESLFAKQVDVQLER----VADEDANL 1004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       103 -NYIK------DSIDTDINFANDVLGYYKILS----EKYKSDLDSIKKYIN----D 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 AVTISVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLE----NNVMTFNVNVKDILNSRF 56
                                                                                                                                                                                                                                                                                                                                                            Kurlandzka A., Rytka J., Gromadka R., Murawski M.;
"A new essential gene located on Saccharomyces cerevisiae chromosome
                                                                                                                                                                                                                                                                                                                                                                                                    Yeast 11:885-890(1995).
-!- FUNCTION: NOT KNOWN; MAY PARTICIPATE IN CELL-CELL AND/OR CELL-SUBSTRATUM INTERACTIONS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              889 LRSVWKPIAAIIGRLNTLVISLAATNETFENINSLFYLKWSACTSLMDIIVAIKIFELKL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       790 DPNNSRDD------YKDLTCSLYEL-----YINKLTILGRDYPIEVDEELLQLFL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EAKVLNYTYEKSN------VEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDY
                                                                                                                                                                                        SEQUENCE FROM N.A.
STRAIN-S288C / AB972;
Barrial B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
Churcher C.M., Connor R., Copsey T., Dear S., Deviin K., Fraser J
Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones I
Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
Walsh S.V., Whitehead S.;
Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
                                                                                                                      Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6.8%; Score 134; DB 1; Length 1150; 20.6%; Pred. No. 2.7; ive 74; Mismatches 134; Indels 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         70 POLY-ASP.
133008 MW; 89688EA09485AC28 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           57 ---NKRENFKNVLESDLIPYKDLTSSNYVVKDPYKFLNK---
                                                   01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                              -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
-!- SIMILARITY: SOME, TO S.POMBE SPAC17H9.20.
                            1150 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---YKTVND----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLY-GLU
                                                                                                                                                                                                                                                                                                                                                MEDLINE=96090137; PubMed=7483852;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; 246881; CAA86966.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 20.6%
Matches 89; Conservative
                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         144 KQGENEKYLPFLNNIETL-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1150 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IRR1.
                                                                                                                                                                                                                                                                                                                                    CHARACTERIZATION.
                                                                                                          IRR1 OR YIL026C.
                                                                                                                                                                 NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S0001288;
                                                                                             IRR1 PROTEIN
                           IRR1_YEAST
P40541;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
RESULT 13
IRR1_YEAST
                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15;
                                                                   1005 NDLPETGFFENIETEFLLFTVKLKGLMK---LNILDERFASRVALNKEKLGPLFKKIVDD 1061
                                                                                                                                                                                   1062 TIMENPEPNKKNIQKAKS----NQTQREKAPLQPNSERETD----HANTENNDPDIPMTID 1114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    [1]

SEQUENCE FROM N.A.

SEQUENCE FROM N.A.

MEDLINE=96337999; PubMed=8688087;

MEDLINE=96337999; PubMed=8688087;

MEDLINE=96337999; PubMed=8688087;

MEDLINE=96337999; PubMed=8688087;

Built C.J., Whithe O., Olsen G.J., Zhou L., Fleischmann R.D.,

Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,

Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,

Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.

Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,

Cotton M.D., Roberts K.M., Hush H.O., Woese C.R., Venter J.C.;

Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;

"Complete genome sequence of the methanogenic archaeon, Methanococcus
-HNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYV---KHFTTPMRKKTM-- 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  214 LMAQKFESERKFKEAAEYYKKSGDTIKEIDEKIAYDEYANSYKWLAIENKYNKEKFEEYI 273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --KQLENNVMTFN-VNVK-----D 50
                                                                                                                                      - - - IQQSSGCFRHLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NK----EKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSEKYKSDLDSIKK----YI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -PYKDLTSSNYVVKDPY-KFL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      101; Indels 102;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 748;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5515ED3B541562EA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6.7%; Score 133; DB 1;
23.0%; Pred. No. 1.9;
iive 38; Mismatches 101;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       748 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           51 ILNSRFNKRENFKNVLE----SDLI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 protein.
748 AA; 90731 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Science 273:1058-1073(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; U67531; AAB98882.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15-JUL-1998 (Rel. 36, Last a HYPOTHETICAL PROTEIN MJ0875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17 EYEVLYLKPLAGVYRSLK--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Methanococcus jannaschii
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                 : || | :| :
1115 LEPIEESSQNNSE 1127
                                                                                                                                                                                                                                                                                    323 AKCTEEDSGSNGK 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 72; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=2190;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TIGR; MJ0875; -
```

```
DARA WANTE FEEFFE
                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license gireement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-ATC. 35210 / B31;
STRAIN-ATC. 35210 / B31;
STRAIN-ATC. 35210 / B31;
BEDLINE-98065943; PubMed-9403685;
Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
Dougherty B., Tomb J.F., Fleischmann R.D., Richardson D.,
Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
Utterback T., Watthey L., McDonald L., Artlach P., Bowman C.,
Garland S., Fujil C., Cotton M.D., Horst K., Roberts K., Hatch B.,
Smith H.O., Venter J.C.;
"Genomic sequence of a Lyme disease spirochaete, Borrelia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     J. Bacteriol. 179:552-556(1997).

-!- FUNCTION: INVOLVED IN THE TRANSMISSION OF SENSORY SIGNALS FROM THE CHEMORECEPTORS TO THE FLACELLAR MOTORS. CHEA IS AUTOPHOSPHORYLATED; IT CAN TRANSFER ITS PHOSPHATE GROUP TO EITHER CHEB OR CHEY (BY SIMILARITY).

-!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).

-!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
           142 NDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLK 201
                                                                                            360 KAKKSLKNVKISNIIFSKYTLECD----ELICRFYLSISQGEFKKSVELLDEYLEISLKI 415
                                 : ::|| |||:|:
|-----TYYQFKYELLNGNYEKA-----LNFLQ
                                                                        202 TIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKF---LSTG-------MVFENLLKSVLSN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ge Y., Charon N.W.;
"An unexpected flaA homolog is present and expressed in Borrelia
burgdorferi.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Trueba G.A., Old I.G., Saint-Girons I., Johnson R.C.; "A cheA cheW operon in Borrelia burgdorferi, the agent of Lyme disease.";
                                                                                                                                                                                                                                                                                                                                                             Borrelia burgdorferi (Lyme disease spirochete).
Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
                                                                                                                                                                                                                                                               044737; P70857; Q44877;
01-NOV-1997 (Rel. 35, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
CHEMOTAXIS PROTEIN CHEM (EC 2.7.3.-).
                                                                                                                                                                                                                                                    864 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN=212;
MEDLINE=98438936; PubMed=9765799;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Res. Microbiol. 148:191-200(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN=212;
MEDLINE=97144545; PubMed=8990312;
                                          323 RSKSFEFAKYMEYLE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 1-30 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Nature 390:580-586(1997).
                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                        LLDWKLARYVKHF 265
                                                                                                                                                                      LSDWKNTRKYKFY 428
                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=139;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             burgdorferi.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KINASES
                                                                                                                                                                                                                                                    CHEA_BORBU
Q44737; P7
                                                                                                                                                                     416
                                                                                                                                        253
                                                                                                                                                                                                                                       CHEA_BORBU
                                         g
                                                                                                                                                                   q
                                                                                                                                                                                                                                                     ద
                                                                                                                                        ò
                                                                          ö
```

```
422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                118 DVLGYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDL--FVIH 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   68 SDLIPYKDL-TSSNYVVKDPYKFLNKEKRDKFLSSYN-----YIKDSIDTDINFAN 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                      (AUTO-) (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8 DNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVLE 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            374 FEIKNVNLESLKS---VRLKE-----DDEAPFKENKNIKKNSPISVNLIRIDSKKID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             318 OMLQALKSLGKIFKSIPDS-SELLADKFFDFVIYYLISNTSE---ESIAKKINLPDVVSH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    176 LEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----YNHN-----NLLTKFLSTGMVFE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               423 YILNLVSEAVISKSSYNQINSEMITLFYNFNYFYDYQESFQRNFLIDLKIVFKDAGLTLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --TPMRK---KTMIQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    64; Mismatches 121; Indels 144;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | | | :: ::| | | :: | | | | 543 PISNIFSRFTRVVRDLSKKLNKIVNLKMEGEETELDKSVIDDLVDPLMHCVRN 595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---CVEN 307
                                                                         InterPro; IPR000410; -.
InterPro; IPR002545; -.
InterPro; IPR002570; -.
Pfam; PF01584; CreW; 1.
Pfam; PF01627; Hpt; 1.
Pfam; PF00512; signal; 1.
PRINTS; PF00344; BCTRLSENSOR.
Sensory transduction; Transferase; Kinase; Phosphorylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 133; DB 1; Length 864; Pred. No. 2.2;
                                                                                                                                                                                                                                                                        L -> I (IN REF. 1).
S -> H (IN REF. 1).
S -> G (IN REF. 1).
L -> S (IN REF. 1).
A -> S (IN REF. 1).
W. C1111DE0877BE624 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QSSGCFRHL-----DEREECKCLLNYKQEGDK-
                                                                                                                                                                                                                                                                                                                                                                                         M.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                            6.78;
EMBL; AE001168; AAC67024.1;
EMBL; U62900; AAC44771.1; --
EMBL; X91907; CAA63002.1; --
TIGR; BB0669; --
                                                                                                                                                                                                                                                51
11
182
182
239
362
559
98352 N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    84; Conservative
                                                                                                                                                                                                                                                                                                                         239
362
559
864 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
                                                                                                                                                                                                                                    Chemotaxis.
MOD_RES
                                                                                                                                                                                                                                                                                       CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 277
```

19;

Search completed: August 8, 2001, 12:32:42 Job time: 210 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

August 8, 2001, 12:30:58 ; Search time 43.63 Seconds Run on:

(without alignments)
658.212 Million cell updates/sec

US-09-500-376-5 Perfect score:

1982 1 AVTTSVIDNILSKIENEYEV......SNFLGISFLLILMLILYSFI 377 Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

219241 segs, 76174552 residues Searched: Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Maximum Match 100% Listing first 45 summaries Post-processing: Minimum Match 0%

PIR 68:\* Database :

pir1:\* pir2:\* pir3:\* pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

•	Description	major merozoite su	probable major sur	major merozoite su		merozoite surface	major merozoite su	major merozoite su	major merozoite su	major merozoite su		major blood-stage	major merozoite su	major merozoite su	merozoj		rhoptry protein -	hypothetical prote	phosphatase (acid	ORF MSV152 probabl	rhoptry protein -	hypothetical prote	ч	hypothetical prote		σ	hypothetical prote	>	Ser/Thr protein ki	hypothetical prote
SUMMAKIES	ΩĬ	SAZQK1	A24594	A45545	S05603	S47282	SAZQGM	A45948	A54498	A26868	A39401	A45604	A28121	A45532	A45546	T18501	T28677	B71612	B71616	T28313	T28676.	A71622	D71607	E71606	G71613	A57719	T18499	T28317	16	T18491
	DB	-	7	7	7	7	H	7	~	~	7	7	7	7	7	7	7	7	7	~	7	7	7	~	ď	ď	7	7	7	7
	Length	1631	1640	400	1639	651	1726	1726	1701	1701	1726	1751	089	1772	1785	3394	2269	3973	2010	1306	2401	507	722	980	1346	706	1619	1127	1714	1119
dР	Query	98.2				58.9	58.7				37.6					•	7.8	•		7.7	٠	•	•	7.5	•		•	•	7.3	7.3
	Score	1947	1942	1824.5	1824.5	1167.5	1164	1157	1145	1131	745	744	588.5	588.5	546	157.5	155.5	155.5	154	153.5	152.5	149	149	149	149	148.5	148.5	147.5	145	144.5
	Result No.	н	7	æ	4	D.	9	7	80	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29

hypothetical prote reticulocyte-bindi repeat organellar asparata over the control of the control	probable membrane hypothetical prote hypothetical prote	
118465 B42771 T18372 T18402 T28301 A64465 T18355 T1835 T09079 T09080 E71622 H64387	C71610 G70163 S67087	
	000	
481 1252 1939 3844 608 1005 1187 2500 2708 2819 1308 377 2523	2013 2166 821	
77.77.77.77.77.77.77.77.77.77.77.77.77.	7.0	
142 142 142 142 141 141 140 140 139 139 139	138.5 138.5 138	
330 330 330 330 441 451	43 44 5	
ŧ		

## ALIGNMENTS

Н	
Ē	_
SULT	C

RESULT 1 SAZQK1 major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)

C; Species: Plasmodium falciparum
C; Species: Plasmodium falciparum
C; Species: Plasmodium falciparum
C; Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 09-Jun-2000
C; Accession: A25120
R; Mackay, M.; Goman, M.; Bone, N.; Hyde, J.E.; Scaife, J.; Certa, U.; Stunnenberg, H. EMBO J. 4, 3823-3829, 1985
A; Title: Polymorphism of the precursor for the major surface antigens of Plasmodium f A; Reference number: A91030; MUID:86136024
A; Rolecule type: DNA
A; Molecule type: DNA
A; Molecule type: DNA
A; Residues: 1-1631 cMAC>
C; Comment: The merozoite stages of different strains have strain-specific surface ant C; Comment: P. falciparum has three stages: sporozoite, merozoite, and gametocyte. The C; Comment: P. falciparum has three stages: sporozoite, merozoite surface antigen
C; Superfamily: major merozoite surface antigen; tandem repeat; transmembrane pr C; Superfamily: major merozoite; surface antigen #status predicted cMAT>
F; 1-19/Domain: signal sequence #status predicted cMBN>
F; 67-84/Region: 3-residue repeats (6-G-T/P)
F; 1614-1631/Domain: membrane anchor #status predicted cMBN>
F; 97, 259, 755, 759, 835, 911, 955, 1049, 1156, 1165, 1436, 1563/Binding site: carbohydrate (Asn

Gaps ; 0 Length 1631; 4; Indels Score 1947; DB 1; Pred. No. 2.8e-95; 2; Mismatches 4; 98.2%; illarity 98.4%; Conservative Query Match Best Local Similarity Matches 371; Conserv

ö

1255 AVTPSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 1314 1 AVTTSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 60 δ g

δ

Dp

121 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV ŏ

1375 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 1434 g

LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 240 181 Q

g

241 VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE 1495 VFENLLKSILSNLLDWKLARYVKHFTTPWRKKTMIQQNSGCFRHLDEREECKCLLNYKQE δλ g ŏ

ij

Gaps

240

us-09-500-376-5.rpr

LGISFLLILMLILYSFI 377

361

```
A;Title: Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene from Pl A;Reference number: S04850; MUID:89345116
A;Accession: S04850

    malaria parasite (Plasmodium falciparum)

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              major merozoite surface antigen precursor - malaria parasite (Plasmodium fal
N;Alternate names: gp195 surface antigen
C;Species: Plasmodium falciparum
C;pate: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jun-2000
C;Accession: S05603; S04850
R;Myler, P.J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1324 NFKNVLESDLIPPKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 1383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1264 AVTPSVIDNILSKIENEXEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 1323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                          NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                   GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 AVTTSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 60
                                                                                                                                        25 AVTPSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDIINSRFNKRE
                                                                                                                                                                                                                                   GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFINNIETLYKTVNDKIDLFVIHLEAKV
                                                                                                                                                                                                                                                                                                                                                                                                                 1 AVTTSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE
                                                                                                                                                                                                                                                                                                                                                                                          LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              241 VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Submitted to the EMBL Data Library, April 1989
A;Reference number: $05603
A;Accession: $05603
A;Accession: $05603
A;Accession: $05603
A;Accession: $05603
A;Cross_references: EMBL:X15063; NID:99896; PIDN:CAA33163.1; PID:99897
R;Myler, P.J.
Nucleic Acids Res. 17, 5401, 1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C;Superfamily: major merozoite surface antigen
C;Keywords: alycoprotein; merozoite; surface antigen
F:1-19/Domain: signal sequence #status predicted <SIG>
F;20-1639/Product: major merozoite surface antigen #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 1824.5; DB 2; Length 1639;
Pred. No. 8.2e-89;
6; Mismatches 17; Indels 1;
                         Length 400;
                       Score 1824.5; DB 2;
Pred. No. 1.7e-89;
                                                                    Mismatches
                                                                 ; 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92.1%;
93.6%;
                       92.1%;
93.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Molecule type: mRNA
A; Residudes: 1504-1639 <MYL2>
A; Cross-treferences: EMBL:X15063
C; Superfamily: major merozoite s
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 93.6'
Matches 353; Conservative
                                                                    Matches 353; Conservative
                                              Similarity
                         Query Match
                                              Best Local
                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                      82
                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                           145
                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                   δ
                                                                                                                                                           g
                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                       òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A45545
major merozoite surface antigen - malaria parasite (Plasmodium falciparum) (fragments)
C;Species: Plasmodium falciparum
C;Species: Plasmodium falciparum
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jun-2000
C;Accession: A45545
R;Blackman, M.J.; Ling, I.T.; Nicholls, S.C.; Holder, A.A.
Mol. Biochem Parasitol. 49, 29-33, 1991
A;Title: Proteolytic processing of the Plasmodium falciparum merozoite surface protein-1
A;Reference number: A45545; MUID:92131048
A;Accession: A45545
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-400 cBLA>
A;Residues: 1-400 cBLA>
A;Note: sequence extracted from NCBI backbone (NCBIN:77612, NCBIP:77621)
C;Superfamily: major merozoite surface antigen; tandem repeat
C;Keywords: glycoprotein; merozoite; surface antigen; tandem repeat
                                                                                                                                                                                                C;Deter: 29-Aug-1987 #sequence_revision 29-Aug-1987 #text_change 09-Jun-2000
C;Accession: A24594
R;Holder, A.A.; Lockyer, M.J.; Odink, K.G.; Sandhu, J.S.; Riveros-Moreno, V.; Nicholls, Nature 317, 270-273, 1985
A;Title: Primary structure of the precursor to the three major surface antigens of Plasm A; Reference number: A24594; MUID:86014355
A;Reference report a 24594
A;Molecule type: DNA
A;Residues: 1-1640 <HOL>
C;Superfamily: major merozoite surface antigen
C;Keywords: surface antigen
                                                                                                                                                        robable major surface antigen (83K, 19K, 42K) precursor - malaria parasite (Plasmodium Species: Plasmodium falciparum ;Date: 29-Aug-1987 #sequence_revision 29-Aug-1987 #text_change 09-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 AVTTSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1324 NFKNVLESDLIPPKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1504 VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1640;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 1942; DB 2;
Pred. No. 5.2e-95;
0; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1615 LGISFLLILMLILYSFI 1631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best_Local Similarity 98.7
Matches 372; Conservative
                                                                                                                                                    probable major
C;Species: Plas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  301
                                          g
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
```

1;

Gaps

us-09-500-376-5.rpr

Wed Aug

g å g ò g ò q ò q

```
Nighternate names: 1978 41970proprotean
C;Species: Plasmodium falciparum
C;Species: Plasmodium falciparum
C;Species: 30-Sep-1987 #sequence_revision 31-Mar-1991 #text_change 09-Jun-2000
C;Accession: A2386; S06361
R;Weber, J.L.; Leaininger, W.M.; Lyon, J.A.
Rwcleic Acids Res. 14, 3311-3323, 1986
A;Title: Variation in the gene encoding a major merozoite surface antigen of the huma A;Reference number: A3386 Mill: 86205236
A;Residues: 1-1104 < WEB1>
A;Cross-references: EMBL:X0831
A;Residues: 1-1104 < WEB1>
A;Cross-references: EMBL:X0831
A;Reference number: S06361; MID:88143999
A;Title: Merozoite surface protein sequence from the Camp strain of the human malaria A;Reference number: S06361; MID:88143999
A;Accession: S06361
A;Accession: S06361
A;Accession: S06361
A;Residues: 1104-1726 < WEB2>
A;Cross-references: EMBL:X03831
C;Comment: The merozoite stages of different strains have strain-specific surface antigen
C;Comment: P: falciparum has three stages: sporozoite, merozoite, and gametocyte. The C;Comment: P: falciparum has three stages: sporozoite, merozoite surface antigen
C;Comment: P: falciparum has three stages: sporozoite, merozoite surface antigen
C;Comment: P: falciparum has three stages: sporozoite, merozoite surface antigen
C;Comment: P: falciparum has three stages: sporozoite, merozoite surface antigen
C;Comment: P: falciparum has three stages: sporozoite, merozoite surface antigen
C;Comment: P: falciparum has three stages: sporozoite, merozoite surface antigen
C;Comment: P: falciparum has three stages: sporozoite, merozoite surface antigen
C;Comment: P: falciparum has three stages: sporozoite, merozoite surface antigen
C;Comment: P: falciparum has three stages: sporozoite, merozoite, merozoite surface antigen
C;Comment: P: falciparum has three stages: sporozoite, merozoite, merozoite surface antigen
C;Comment: P: falciparum has three stages: sporozoite, merozoite, merozoite surface antigen
C;Comment: P: falciparum has three stages of falciparum has three stages of falciparum has three stages of falciparum 
                                                                                                               (Plasmodium falciparum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ä;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7 IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |||||: :| ::||: ||: ||: || || ||: : ||: :||: ||: :||: ||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :|: :|: :|: :|: :|: :|: :|: :||: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEKYKSDLDSIKKYI--------NDKQGENEKYLPFLNNIETLYKTVN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DKIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NHNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 1726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels

    malaria parasite

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     58.7%; Score 1164; DB 1;
57.5%; Pred. No. 5.9e-54;
iive 60; Mismatches 84;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1726
                                                                                                               major merozoite surface antigen precursor N; Alternate names: 195K glycoprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           YPLFDGIFCSSSNFLGISFLLILMLILYSFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         YPLFDGIFCSSSNFLGISFLLILMLILYSFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 225; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         347
                                                                        SAZQGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A45948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           म
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        merc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RO-71)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C-terminal region of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Metrozoite surface antigen 1 - malaria parasite (Plasmodium falciparum) (strain R C; Species: Plasmodium falciparum A; Variety: strain R0-71
C; Species: Plasmodium falciparum A; Variety: strain R0-71
C; Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 09-Jun-2000
C; Accession: 547282
R; Tolle, R; Bujard, H; Cooper, J.A.
submitted to the EMBL Data Library, July 1994
A; Bescription: Plasmodium falciparum: recombination within the C-terminal region A; Reference number: 547282
A; Reference number: 547282
A; Residues: 1-651 <TOL>
A; Cross-references: EMBL: 235329; NID: 9535257; PIDN: CA884558.1; PID: 9535258
A; Experimental source: strain R0-71
C; Superfamily: major merozoite surface antigen
C; Keywords: glycoprotein; merozoite; surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (strain
                                                                                                                                                                                1563 GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSNF 1622
                                                                                                                                                                                                                                                                                               300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LAKYKDDLESIKKVIKEEKEFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLVN 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCY 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEKYKSDLDSIKKYI--------NDKQGENEKYLPFLNNIETLYKTVND 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYN 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        228 HNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDE 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     7 IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL 66
GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV
                                      VFENILKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE
                                                                                                                                             LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                                                                                                                         GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    58.9%; Score 1167.5; DB 2; Length 651; 57.9%; Pred. No. 1.3e-54; ive 59; Mismatches 84; Indels 21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PLFDGIFCSSSNFLGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1623 LGISFLLILMLILYSFI 1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 57.9
Matches 226; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                    1384
                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       361
121
                                                                                                                                                                                                                                                                                       241
                                                                                                                                                                                                                                                                                                                                                                                                                                         301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
S47282
```

ò g ò q δ g ò g ð g ð

셤

ð 셤 ò g à Q ò 셤 ö a ð g ò

```
C;Accession: A26868
R;Tanabe, K.; Mackay, M.; Goman, M.; Scaife, J.G.
J. Mol. Biol. 195, 273-287, 1987
A;Title: Allelic dimorphism in a surface antigen gene of the malaria parasite Plasmod A;Reference number: A26868; MUID:88011243
A;Accession: A26868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 09-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1493 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD 1552
                                                                                                                                                                                                                                                                                                                                                              1313 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNITTFNLNLNDILNSRLKKRKYFLDVL 1372
                                                                                                                                      1609 LDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKP 1668
                                                                                                                                                                                                                                                                                                                                             EREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDC 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL
                                                                                                                    ----NDKQGENEKYLPFLNNIETLYKTVN
                                                                                                                                                                                                                                                                   NHNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLD
                                                                                                                                                                                                                                                                                                        1553 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKF-QDMLNISQHQCVKKQCPENSGCFRHLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEKYKSDLDSIKKYI-------NDKQGENEKYLPFLNNIETLYKTVN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NHNNLLTKFLSTGMY--FENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DKIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDY
                                                                                                                                                                                            DKIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Keywords: surface antigen
:1-19/Domain: signal sequence #status predicted <SIG>
:20-1701/Product: major merozoite surface antigen #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1701;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   57.1%; Score 1131; DB 2;
56.5%; Pred. No. 3.2e-52;
iive 60; Mismatches 85;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A,Molecule type: DNA
A,Residues: 1-1701 <TANN
C,Superfamily: major merozoite surface antigen
C,Keywords: surface antigen
C,1-19/Domain: signal sequence #status predicte
F;20-1701/Product: major merozoite surface anti
                                                                                                                                                                                                                                                                                                                                                                                                                                        1671 YPLFDGIFCSSSNFLGISFLLILMLILYSFI 1701
                                                                                                                                                                                                                                                                                                                                                                                                                     347 YPLFDGIFCSSSNFLGISFLLILMLILYSFI 377
                                                                                                                    SEKYKSDLDSIKKYI-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 57.1%
Best Local Similarity 56.5%
Matches 222; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1493
                                                                               1373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         285
                                          67
                                                                                                                    127
                                                                                                                                                                                                                                                                     227
                                                                                                                                                                                                                                                                                                                                               287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   227
                                                                                                                                                                                            167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A26868
                                                                             g
                                                                                                                                                     a
                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                        ò
                                                                                                                  δ
                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C; Accession: A54498

Well 21, Coppel, R.L.; McIntyre, P.; Langford, C.J.; Woodrow, G.; Brown, G.V.; Mol. Blochem. Parasitol. 27, 291-302, 1988

A; Title: Variation in the precursor to the major merozoite surface antigens of Plasmodiu A; Reference number: A54498; MUD:88142999

A; Status: preliminary
                                                                                 meroz
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKV 1457
                                                                                                                                                                                                                                                                                                                                                                                                                                        ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                286
                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Molecule type: DNA
A; Residues: 1-1701 <PET>
A; Residues: 1-1701 <PET>
Cross-references: GB:M19143; NID:g160412; PIDN:AAA29653.1; PID:g160413
C; Superfamily: major merozoite surface antigen
C; Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                7 IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----NDKQGENEKYLPFLNNIETLYKTVN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NHNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DKIDLEVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1701;
                                                                                                                                                                                                                                                                                        Length 1726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   57.8%; Score 1145; DB 2;
56.8%; Pred. No. 5.8e-53;
ive 61; Mismatches 86;
                                                                                                                                                                                                                                                                                        58.4%; Score 1157; DB 2; 57.3%; Pred. No. 1.4e-53;
                                                                                                                                                                                                                                                                                                                           60; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1696 YPLFDGIFCSSNFLGISFLLILMLILYSFI 1726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 YPLFDGIFCSSSNFLGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEKYKSDLDSIKKYI------
                                                                                                                                                                                                                                                                                                          Best Local Similary
Matches 224; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity
Matches 222; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ^
                                                                                                                                                                                                                                                                                                                                                                                                                                      67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
```

4

ô

us-09-500-376-5.rpr

9

Gaps

48;

Indels

Length 1751;

```
1604 VKSSGLIEKLMKSKLIKENESKEILSQLLNVQTQLLTMSSEHTCI----DTNVPDNAACY 1659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C;Species: Plasmodium yoelii
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 26-Aug-1999
                                                                                                                                                                                                                                                    LTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSEKYKSDLD 135
                                                                                                                                                                                                                                                                                                                              ----KQGENEKYLPFLNNIETLYKTV 165
                                                                                                                                                                                                                                                                                                                                                                                                         166 NDKIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTD 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KYKSDLDSIKKYI-----ND-KQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | : | ::|: | :| | :| | 429 VYNTQLAAVKEQIATIEAETNDTNKEEKKKYIPILEDLKGLYETVIGQAEEYSEELQNRL 488
(NCBIN:83591, NCBIP:83592)
                                                                                                                                                                        16 NEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVLESDLIPYKD
                                                                                                                                                                                                                                                                                                                                                                                                                                             1551 VSKVNTYTDNLKKVINNCQLEKKEAEITVKKLQDYNKMDEKLEEYKK-----SEKKNE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RHLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NILSKI ENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDI LNSRFNKRENFKNVLES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   226 YNHNNLLTKFLSTGMVFENLLKSVLSNLLDWK---LARYVKHFTTPMRKKTMIQQSSGCF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  major merozoite surface antigen – Plasmodium yoelii (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Note: the authors translated the codon GTA for residue 429 C;Superfamily: major merozoite surface antigen C;Reywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      R.Burns Jr., J.M.; Daly, T.M.; Vaidya, A.B.; Long, C.A.
Proc. Natl. Acad. Sci. U.S.A. 85, 602-606, 1988
A.;Title: The 3' portion of the gene for a Plasmodium yoelii
A;Reference number: A28121; MUID:88124889
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Residues: 1-680 <BUR>
A;Cross-references: GB:J03612; NID:g160678; PID:g160679
A;Experimental source: strain 17XL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 2;
                                                                                            37.5%; Score 744; DB 2; I
38.9%; Pred. No. 8.4e-32;
ive 76; Mismatches 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; Score 588.5; DB 2;
; Pred. No. 4.6e-24;
79; Mismatches 142;
A:Note: sequence extracted from NCBI backbone c:Superfamily: major merozoite surface antigen C:Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               343 KPDCYPLFDGIFCSSSNFLGISFLLILMLIL 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             29.78;
34.48;
                                                                                                                                                                                                                                                                                                                              SIK ---KYIND-----
                                                                                                              Best Local Similarity 38.9%
Matches 152; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Accession: A28121
A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C; Accession: A28121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Simi
Matches 132;
                                                                                            Query Match
Best Local &
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               283
                                                                                                                                                                                                                                                  9/
                                                                                                                                                                                                                                                                                                                              136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1660
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         369
                                                                                                                                                                          ò
                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                  Óγ
                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                          merozoite surface antigen 1 precursor - Plasmodium vivax
C.Species: Plasmodium vivax
C.Species: Plasmodium vivax
C.Species: Preb-1992 #sequence_revision 28-Feb-1992 #text_change 24-Nov-1999
C.Accession: A39401
R.del Portillo, H.A.; Longacre, S.; Khouri, E.; David, P.H.
Proc. Natl. Acad. Sci. US.A. 88, 4030-4034, 1991
A.Title: Primary structure of the merozoite surface antigen 1 of Plasmodium vivax reveal
A.Reference number: A39401
A.Accession: A39401
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-1756 <-DEL>
A.Cross-references: GB:M60807
C.Superfamily: major merozoite surface antigen
C.Seywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1466 AVNEEVKKVEADIKAEDDKIKKIGSDSTKTTEKTQSMAKKAELEKYLPFLNSLQKEYESL 1525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1526 VSKVNTYTDNLKKVINNCQLEKKEAEITVKKLQDYNKMDEKLEEYKK------SEKKNE 1578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1579 VKSSGLLEKLMKSKLIKENESKEILSQLLNVQTQLLTMSSEHTCI----DTNVPDNAACY 1634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ....---DSIKKYIND------KQGENEKYLPFLNNIETLYKTV 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              166 NDKIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTD 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RHLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECT 342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         76 LTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSEKYKSDL- 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVLESDLIPYKD 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        226 YNHNNLLTKFLSTGMVFENLLKSVLSNLLDWK---LARYVKHFTTPMRKKTMIQQSSGCF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 37.6%; Score 745; DB 2; Length 17 Best Local Similarity 39.4%; Pred. No. 7.3e-32; Matches 154; Conservative 72; Mismatches 117; Indels
                1691 KEGSEPLFEGVFCSSSFLSLSFLLLMLLFL 1721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              343 KPDCYPLFDGIFCSSSNFLGISFLLILMLIL 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A45604
                ŏ
                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

merozoite surface antigen

1;

31;

680;

Length Indels

Len

as

ω

Aug

Wed

ò g ò g õ g ò g

```
major merozoite surface antigen precursor - Plasmodium chabaudi chabaudi
C; Species: Plasmodium chabaudi chabaudi
C; Species: 30-8ep-1993 #sequence_revision 30-8ep-1993 #text_change 26-Aug-1999
C; Accession: A45546
R; Delecrinider, W.; Hendrix, D.; Bendahman, N.; Hanegreefs, J.; Brijs, L.; Hamers-Ca A, Tichem. Parasitol. 43, 231-244, 1990
A,Title Molecular cloning and sequence analysis of the gene encoding the major meroz A, Reference number: A45546
A; Steperence number: A5546
A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Cross_references: EMBL:Z98551; NID:e1331903; PID:e1331922; PIDN:CAB11140.1
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C; Species: Plasmodium falciparum
C; Date: 15-0ct-1999 #sequence_revision 15-0ct-1999 #text_change 09-Jun-2000
C; Accession: T18501
R; Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A; Reference number: 218935
A; Accession: T18501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1469 LDVLSYELALFKDINTNKFVVKNPYQLLDNDKKDKQMINLKYAIKGVTEDIETATDGIEF 1528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1589 HKLENYKIEKAGFDILMANLETYIRIDEKLEDFVESAEKNKHIASIA------LN 1637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1638 NLNKSGLVTEGESKKILAKMLNMDAMDLLGIGSNHVCISTSTPDNAGCFRYDDGTEEWRC 1697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          63 KNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGY 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        123 YKILSEKYKSDLDSIKKYI----NDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLE 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 TTSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENF 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        178 AKVLNYTYEKSNVEVKIKELNYLKTIQDKLADF----KKNNNFVGIADLSTDYNHNNLLT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         234 KFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LLNYKQ - - EGDKCVENPNPTCNENNGGCDADAKCTEEDSGSN - - GKKITCECTKPDCYPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hypothetical protein C0760c - malaria parasite (Plasmodium falciparum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 1785;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PID:9160598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           160;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 546; DB 2;
Pred. No. 2.4e-21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Residues: Î-1785 <DEL>
A;Cross-references: GB:N34947; NID:g160597; PII
Csuperfamily: major merozoite surface antigen
C;Reywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              : |:||||| |:|:| ||| |:|::
XAGVFCSSSGFMGLSILLIITLIVFN 1783
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       350 FDGIFCSSSNFLGISFLLILMLILYS 375
                              1747 GVFCSSSSFMGLSILLIITLIVFN 1770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               27.5%; Scot.
30.8%; Pred
83; W
     352 GIFCSSSNFLGISFLLILMLILYS 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 119; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Molecule type: DNA
A; Residues: 1-3394 < LAW>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: mRNA
                                                                                                                            14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       294
     ò
                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C. Accession: A45532, A45531

R. Lewis, A. P.

Mol. Blochem. Parasitol. 36, 271-282, 1989

A. Title: Cloning and analysis of the gene encoding the 230-kilodalton merozoite surface A: Reference number: A45532; MUD: 90014981

A. Reference number: A45532; MUD: 90014981

A. Accession: A45532

A. Status: preliminary

A. Molecule type: DNA

A. Residues: 1-1772 < LEWA

A. Residues: 1-1772 < LEWA

A. Cross-references: GB: 304668; NID: 9160492; PID: 9160493

R. Daly, T. M.; Burns Jr., J. M.; Long, C. A.

Mol. Blochem. Parasitol. 36, 283-285, 1989

A. Title: Precursor to the major merozoite surface antigen of Plasmodium yoelii: cloning
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            major merozoite surface antigen precursor - Plasmodium yoelii
C;Species: Plasmodium yoelii
C;Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 26-Aug-1999
C;Accession: A45532; A45531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          :| :|||:::||:::||: |:::|| :: |
1461 ELDLFKDLSTNKYVIRNPYQLLDNDKKDKQIVNLKYATKGINEDIETTTDGIKFFNKAVE 1520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1630 KSGLVGEGESKKILAKMLNMDGMDLLGVDPKHVCVDTRD---IPKNAGCFRDDNGTEEWR 1686
                                                                                                292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KYKSDLDSIKKYI-----ND-KQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      351
  ---KKNNNFVGIADLSTDYNHNNLLTKFL 236
                                                                                                                                              594
                                                                                                                                                                                                 351
                                                                                                                                                                                                                                                 654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVLES 68
                              | | | | : | : | | | | : | : | | DNYKNEKAEFEILTKNLEKYIQIDEKLDEFVEHAENNKHIASIA------LINNLN
                                                                                                                                 CLLNYKQ-EGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFD
                                                                                                                                                                                                                       LNYTYEKSNVEVKIKELNYLKTIQDKLADF----KKNNNFVGIADLSTDYNHNNLLTKFL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CLLNYKQ-EGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFD
                                                                                                STGMVFENLLKSVLSNLLDWK - - - - LARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STGMVFENLLKSVLSNLLDWK - - - - LARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 1772;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Accession: A45531
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 454-1094 ADAL>
A; Cross-references: GB:J03975; NID:g160081; PID:g160082
C; Superfamily: major merozoite surface antigen
C; Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  34.4%; Pred. No. 1.1.
ive 79; Mismatches 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                29.7%; Score 588.5;
LNYTYEKSNVEVKIKELNYLKTIQDKLADF-
                                                                                                                                                                                                                                                                                                                     GIFCSSSNFLGISFLLILMLILYS 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              132; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best_Local Similarity
Matches 132; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                             489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          129
  181
                                                                                              237
                                                                                                                                                                                                 293
                                                                                                                                                                                                                                                                                               352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      293
```

δ ద å g õ g ò g δ a

A; Map position: 3 A; Note: C0760c

```
22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |: |: |: || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || |- || 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   :: | :: | : | | 1481 | 1481 | 1481 | 1482 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 1483 | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                55 RFNKRENFKNVL-ESDLI---PYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSID 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     111 TDINFAN-DVLGYYKILSEKYKSDLDSI-----KKYINDK----QGENEKYLPF---- 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  210 -----FKKNNNFVGIAD-----LSTDYNHNNLLTKFLSTGMVFENLLKSVLSNL 253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             254 LDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQEGDKCVENPNPTCN 313
Query Match 7.9%; Score 157.5; DB 2; Length 3394;
Best Local Similarity 24.8%; Pred. No. 1.5;
Matches 90; Conservative 61; Mismatches 121; Indels 91; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                      8 DNILSKIENEYEVLYLKPLAGVYRSLK------KQLENNVMTFNVNYKDILNS 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
```

Search completed: August 8, 2001, 12:31:02 Job time: 110 sec

1524 INN 1526

314 ENN 316

```
NAME/KEY: Modified-site
                    1031
1031
2710
2710
990
990
2366
2366
2366
903
903
778
720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                            US-08-290-919-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CENGTH:
112.5
1112.5
1112.5
1112.5
1112.5
1100.5
1100.5
100.5
100.5
100.5
100.5
100.5
100.5
100.5
100.5
100.5
100.5
100.5
100.5
100.5
August 8, 2001, 12:30:01; Search time 34.41 Seconds (without alignments) 225.590 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 120, App
Sequence 120, App
Sequence 4, Appli
Sequence 4, Appli
Sequence 186, App
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 3, Appli
                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 2, Apl
Sequence 2, Apl
Sequence 2, Apl
Sequence 189, Apl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 11,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 12,
Sequence 2, 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Description
                                                                                                                                       1 AVTTSVIDNILSKIENEYEV.....SNFLGISFLLILMLILYSFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                          /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
          GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         S-08-257-073-3
S-08-184-009-120
S-08-458-356-120
S-08-568-459A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3-08-755-587-186
3-08-415-593-43
3-08-392-625-21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -09-285-502-2
-08-974-549A-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -08-645-193B-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S-08-290-919-11
S-08-290-919-1
S-08-290-919-12
S-08-290-919-2
S-08-446-855A-2
                                                                                                                                                                                                                 Potal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5-09-150-741-2
5-08-973-462-8
5-08-323-170B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                08-466-961A-21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-487-826B-4
                                                                                                                                                                                             197339 seqs, 20590346 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
                                                                                                                                                                                                                                                                     Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                  OM protein – protein search, using sw model
                                                                                                                                                             BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                 Issued_Patents_AA:*
                                                                                                                                                                                                                                      Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                   US-09-500-376-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              B
                                                                                                                                                                                                                                                                                                                                                                                                                                                         %
Query
Match Length [
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         455
11147
11147
11147
11239
1239
1239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2391
2391
1786
3135
984
984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                984
1435
1435
1494
978
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      455
                                                                                                                                                                                                                                                                       Post-processing:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          116.5
116.5
113.5
113.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     124
122
121.5
121.5
121.5
119.5
119.5
                                                                                                                               Perfect score:
                                                                                                                                                             Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      198
125.5
125.5
                                                                                                                                                                                                                                                                                                                 Database :
                                                                                                                                         Sequence:
                                                                                                                                                                                              Searched:
                                                                         Run on:
                                                                                                                  Title:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

```
28 112.5 5.7 1007 3 US-08-851-843A-86 Sequence 86, Appl 29 112.5 5.7 1007 4 US-08-974-549A-187 Sequence 187, App 30 112.5 5.7 1031 3 US-08-851-843A-2 Sequence 1.0 App 31 112.5 5.7 1031 3 US-08-974-549A-110 Sequence 2, Appli 32 110.5 5.6 2710 1 US-08-480-604A-6 Sequence 6, Appli 33 110.5 5.6 2710 2 US-08-975-2625-20 Sequence 6, Appli 34 109.5 5.5 990 2 US-08-932-625-20 Sequence 20, Appli 35 109.5 5.5 990 2 US-08-645-193B-15 Sequence 15, Appli 36 109.5 5.5 990 2 US-08-645-193B-15 Sequence 15, Appli 37 109.5 5.5 930 2 US-08-645-193B-15 Sequence 10, Appli 38 109.5 5.5 930 2 US-08-95-12-24 Sequence 10, Appli 39 108.5 5.5 903 4 US-08-95-13-14 Sequence 11, Appli 42 108 5.4 439 2 US-08-853-659A-7 Sequence 37, Appli 43 107 5.4 778 5 PCT-US93-03076-4 Sequence 24, Appli 44 106.5 5.4 720 2 US-08-910-551B-2 Sequence 2, Appli 45 106.5 5.4 720 2 US-08-910-551B-2 Sequence 1, Appli 45 106.5 5.4 720 2 US-08-940-236-1
```

## ALIGNMENTS

```
Sequence 3, Application US/08290919
Patent No. 5720959
CORRESPONDENCE ADDRESS:
ADDRESSES: 100 NEW YORK AVENUE, N.W.
TITLE OF INVENTION: URPROVEMENTS IN OR RELATING TO MALARIA NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSES: 100 NEW YORK AVENUE, N.W.
ITHE OF INVENTION: US.A.
2TRET: 1100 NEW YORK AVENUE, N.W.
2TRET: WASHINGTON NEW YORK AVENUE, N.W.
COMPUTER: IBM PC COMPATIABLE PORM:
MEDIUM TYPE: PloPPY disk
COMPUTER: IBM PC COMPATIABLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CLASSIFICATION DATA:
APPLICATION NUMBER: US/08/290.919
FILING DATE: 22-FEB-1992
PRIOR APPLICATION NUMBER: 22-FEB-1993
ATTORIEX APPLICATION NUMBER: 21-FEB-1993
ATTORIEX APPLICATION NUMBER: 21-FEB-1993
ATTORIEX COMMUTICATION INFORMATION:
RESTRANCE CHARACTERISTICS:
LEEFENA: (202) #22-0944
TELEPHONE: 2000
TELEFAX: (202) #22-0944
TELENGH: 53 amino acid
STRANDEDNESS: Single
PRATURE: PROCEET NUMBER: 21-300
TELEFAX: (102) #21-300
```

ö

Gaps

ö

```
263 KHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQ-EGDKCVENPNPTCNENNGGCDA 321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12 KHVCVDTRD---IPKNAGCFRDDNGTEEWRCLLGYKKGEGNTCVENNNPTCDINNGGCDP 68
                                              308 PNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSN 359
                                                                                                                                                                                           Sequence 11, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
APPLICANT: HOLDEN, ANTHONY A.
APPLICANT: BLACKMAN, MICHARL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENITON: IMPROVEMENTS IN OR RELATING TO MALARIA TITLE OF INVENITOS: 19
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
                                                                     2 PNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPDSYPLFDGIFCSSSN
  3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919
FILING DATE: 04-0CT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 16,773
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12.3%; Score 243; DB 1;
48.0%; Pred. No. 5.4e-12;
tive 14; Mismatches 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               69 TASCQNAESTENSKKIICTCKEPTPNAYYEGVFCSSSS 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        322 DAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSN 359
                                                                                                                                                                                                                                                                                                                                                                                                                              CUSHMAN DARBY & CUSHMAN, L.L.P.
  Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                   STREET: 1100 NEW YORK AVENUE, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 106 amino acids
TYPE: amino acid
STRANDEDNESS: single
  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WASHINGTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
Matches 47; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STATE: D.C
                                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY:
  48;
                                                                                                                                                                                US-08-290-919-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-290-919-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 4
US-08-290-919-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
  Matches
                                                ò
                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                ö
                                                                                                                                                                                  Gaps
                                                                                                                                                                                ö
                                                                                                                                                                                                                       308 PNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSN 359
                                                                                                                                                                                                                                                 2 PNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSN 53
                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
                                                                                                                                    Length 53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 53;
                                                                                                                                                                             1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY: U.S.A.

ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUW TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919
FILLING DATE:
                    /label= X /note= "X = M and N, or N"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 212242/HCM/MJL/6BC8,
                                                                                                                               Score 296; DB 1;
Pred. No. 1.8e-16;
0; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 277; DB 1;
Pred. No. 5.3e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P STREET: 1100 NEW YORK AVENUE, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /label= X /note= "X = M and N,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 6B 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INPORMATION.
NAME: KOKULIS, PAUL N.
                                                                                                                                                                                                                                                                                                                                                                                  Sequence 4, Application US/08290919 Patent No. 5720959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 21
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
                                                                                                                               14.9%;
98.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                14.0%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME/KEY: Modified-site
                                                                                                                             Query Match
Best Local Similarity 98.19
Matches 51; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 53 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WOLECULE TYPE: protein
             CTHER INFORMATION:
CTHER INFORMATION:
US-08-290-919-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; OTHER INFORMATION:
; OTHER INFORMATION:
US-08-290-919-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WASHINGTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRANDEDNESS:
LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LOCATION:
                                                                                                                                                                                                                                                                                                                                                         US-08-290-919-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CITY: W
                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                         õ
```

2

Length 106;

```
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CITY: WASHINGTON
                                                       WASHINGTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STATE: D.C
                                                                      STATE: D.C
COUNTRY: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE:
                             ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-290-919-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
TITLE OF INVENTION: VACCINE
                        GENERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 201; DB 1; Length 48
Pred. No. 3.3e-09;
1; Mismatches 4; Indels
                                                                                                                                                                           272 KTMIQOSSGCFRHLDEREECKCLLNYKQEGDKCVENPNPT 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ) OTHER INFORMATION: /label= X ; OTHER INFORMATION: /note= "X = M and N, Or N" US-08-290-919-1
                                                                                                                                             STREET: 1100 NEW YORK AVENUE, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 12, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
Sequence 1, Application US/08290919
Patent No. 5720959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match 10.1%;
Best Local Similarity 87.5%;
Matches 35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME/KEY: Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               single
                                                                                                                                                            CITY: WASHINGTON STATE: D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-290-919-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 2, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 201; DB 1; Length 48;
Pred. No. 3.3e-09;
1; Mismatches 4; Indels
                                                                                                                                                                                                                 CURREATION SISTEM:
CURRENT APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION DATA:
FILING DATE:
APPLICATION NUMBER:
CLASSIFICATION NUMBER:
BY CLASSIFICATION NUMBER:
APPLICATION NUMBER:
APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE:
APPLICATION NUMBER:
FILING DATE:
APPLICATION NUMBER:
FILING DATE:
TILING DATE:
APPLICATION NUMBER:
NAME:
REGISTRATION NUMBER:
16,773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          272 KTMIQQSSGCFRHLDEREECKCLLNYKQEGDKCVENPNPT 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 212242/HCM/MJL/6BC8/
CUSHMAN, L.L.P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CUSHMAN DARBY & CUSHMAN, L.L.P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STREET: 1100 NEW YORK AVENUE, N.W
                1100 NEW YORK AVENUE, N.W
                                                                                   COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE/DOCKET NUMBER: 21
TELECOMMUNICATION INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEPHONE: (202) 861-3000
TELEFAX: (202) 861-3094
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 12:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10.1%;
87.5%;
CUSHMAN DARBY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 10.15
Best Local Similarity 87.55
Matches 35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 48 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; MOLECULE TYPE: peptide US-08-290-919-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CORRESPONDENCE ADDRESS:
```

ö

```
GENERAL INFORMATION:

APPLICANT: Stewart et al.

TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate

Patent No. 6183996

TITLE OF INVENTION: Synthetase II

FILE REPERENCE:

CURRENT FILING DATE: 1998-09-10

EARLIER PAPLICATION NUMBER: D16380

EARLIER PAPLICATION NUMBER: P16380

EARLIER PILING DATE: 1993-12-02

EARLIER APPLICATION NUMBER: 08/46,855

EARLIER PILING DATE: 1993-12-02

EARLIER PILING DATE: 1995-07-06

NUMBER OF SEQ ID NOS: 15

SEQ ID NO 2

LENGTH: 2391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1392 DINILSAQNSNNNFSCNNENMNKANVDVNVLENDT------KKREDINTTTVFMEGQ 1442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1532 -----YLDNFNTSDEEIGNNKNMDMYLSKEKSISNKNPGNSYYVVDSVYNNEYK 1580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1443 NSVINNKNKENSSLLK------GDEEDIVMVNLK-KENNYNSVINNVDCRKKDM 1489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             86 PYKFLNKEKRDKFLSSYNYIKDSIDTDIN-FANDVLGYYKILSEKYKSDLDSIKKY---- 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   141 ---INDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKVLNYTYEKSNVEVKIKEL 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NYLKTIQDKLADFKK-----NNNFVGIADLSTDYNHNNLLTKFLSTGMVFENLLKSV 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---MRKKTMIQQSSG-----CFRHLDE 287
                                                                                                                                                                                                                                                                                                                                    Query Match 6.3%; Score 125.5; DB 2; Length 2391; Best Local Similarity 19.3%; Pred. No. 0.37; Matches 68; Conservative 62; Mismatches 111; Indels 111; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 125.5; DB 4; Length 2391; Pred. No. 0.37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; OTHER INFORMATION: Description of Artificial Sequence: protein US-09-150-741-2
                                                                                                                                                                                                                                                                                                                                                                                                                                          37 LENNV---MTFNVNVKDILNSRFNKRENFKNVLESDLIPYKDLTSSNYVV-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      : | | : | | : | | 1490 DG-KNINDECKTYKKNKYKDMGLNNNIVDELSNGTSHSTNDHL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             288 REECKCLLNYKQEGDKCVENPNPTC-NENNGGCDADAK---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2, Application US/09150741 Patent No. 6183996
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM: Artificial Sequence
                                                TELEFAX: 703-816-4100 INFORMATION FOR SEQ ID NO: 2:
                                                                                                SEQUENCE CHARACTERISTICS:
LENGTH: 2391 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         250 LSNLLDWKLARYVKHFTTP---
          TELEPHONE: 703-616-4100
                                                                                                                                                      TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                           , MOLECULE TYPE: protein US-08-446-855A-2
                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-150-741-2
                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                             δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: Flores, Maria V
APPLICANT: O'SULLIVAN, William J
TITLE OF INVENTION: Phosphate synthetase II
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderhye PC
STREET: ILOO NO. 5849573th Glebe Road, 8th Floor
CITY: Arlington
STATE: Virgina
COUNTRY: USA
ZIP: 222014714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4; Indels
        CURENT AFFILATION DAIRS:
APPLICATION UNBER: US/08/290,919
FILING DATE: 04-0CT-1994
CLASSIFICATION 435
PRIOR APPLICATION DATA:
APPLICATION NUBER: GB 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATTORREY/ABERT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 16,773
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEFAX: (202) 822-0944
TELEFAX: (202) 822-0944
INFORMATION FOR SCO ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       272 KTMIQQSSGCFRHLDEREECKCLLNYKQEGDKCVENPNPT 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9 KKQCPENSGCFRHLDEREECKCLLNYKQEGDKCVENPNPT 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10.0%; Score 198; DB 1;
85.0%; Pred. No. 5.6e-09;
tive 2; Mismatches 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; LOCATION: 1
; OTHER INFORMATION: /label- X
; OTHER INFORMATION: /note= "x = M and N, or US-08-290-919-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIOLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NATA:
APPLICATION NUMBER: US/08/446,855A
FILING DATE: 06-Jul-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2, Application US/08446855A Patent No. 5849573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE/DOCKET NUMBER: 47-80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29.009
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME/KEY: Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 10.03
Best Local Similarity 85.03
Matches 34; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: amino acid
STRANDEDNESS: si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-446-855A-2
```

õ g S

```
qq
        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           17;
    15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: DAUBERSIES, PIERRE
APPLICANT: DAUBERSIES, PIERRE
TITLE OF INVENTION: MALARIAL PERFERENCYTIC STAGE POLYPEPTIDE MOLECULES
TITLE OF INVENTION: MALARIAL PERFERENCE: 0660-0125-0 PCT
CURRENT FILING DATE: 1996-02-06
EARLIER APPLICATION NUMBER: US/O8/973,462B
EARLIER APPLICATION NUMBER: US/O8/973,462B
EARLIER PILING DATE: 1996-06-10
EARLIER FILING DATE: 1996-06-12
EARLIER FILING DATE: 1995-06-13
NUMBER OF SEQ ID NOS: 29
                                                                                                                                                                                                                                                                                                                            DTNILSAQNSNNNFSCNNENMNKANVDVNVLENDT------KKRREDINTTTVFMEGQ 1442
                                                                                                                                                                                                                                                                                                                                                                                                                                                          1532 -----YLDNFNTSDEEIGNNKNMDMYLSKEKSISNKNPGNSYYVVDSVYNNEYK 1580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     :|||: :: | |:| :| | :| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| |
                                                                                                                                                                                                                                       86 PYKFLNKEKRDKFLSSYNYIKDSIDTDIN-FANDVLGYYKILSEKYKSDLDSIKKY---- 140
                                                                                                                                                                                                                                                                                                                                                                                                            ---INDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKVLNYTYEKSNVEVKIKEL 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NYLKTIQDKLADFKK-----NNNFVGIADLSTDYNHNNLLTKFLSTGMVFENLLKSV 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---FLNKEKR----DKFLSSYNYIKDSIDTDINFANDVLGYYKILSEKYKSDLDSIKKYI- 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               142 -NDKQGENEKYLPFLNNIETLYK---TVNDKIDLFVIHLEAKVLNYTYEKSNVEVK-IKE 196
    Indels 111; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 VTTSVIDNI-----LSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNV---KD- 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          250 LSNLLDWKLARYVKHFTTP-------MRKKTMIQQSSG------CFRHLDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   51 ···ILNSR------FNKRENFKNVLESDLIPYKDLTSSNYVVKDPYK-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1490 DG-KNINDECKTYKKNKYKDMGLNNNIVDELSNGTSHSTNDHL-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      OTHER INFORMATION: Description of Artificial Sequence: Polypeptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              288 REECKCLLNYKQEGDKCVENPNPTC-NENNGGCDADAK------CTEED 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 4; Length 1786;
                                                                           37 LENNY---MTFNVNVKDILNSRFNKRENFKNVLESDLIPYKDLTSSNYVV--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 6.3%; Score 124; DB 4; Le
Best Local Similarity 25.3%; Pred. No. 0.33;
Matches 68; Conservative 45; Mismatches 80;
Mismatches 111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 8, Application US/08973462B Patent No. 6191270
62;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: Patentin Ver. 2.0 SEQ ID NO 8
    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1786
68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 9
US-08-973-462-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-973-462-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    198
    Matches
                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
```

```
GENERAL INFORMATION:

GENERAL INFORMATION:

APPLICANT: Williamson, Kim C.

APPLICANT: Waslow, David C.

APPLICANT: Waslow, David C.

TITLE OF INVENTION: Clouding and Expression of Plasmodium

TITLE OF INVENTION: faliciparum Transmission-Blocking Target Antigen, Pfs230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----- AN 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DVLGYYKILSEKYKSDLDSI---KKYIND------KQGENEKYLPFLNNIETLYKTVND 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               168 KIDLEVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYN 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----NKDYE 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10 LFLIILVYIF-SEKKDLRCNVIKGN-NIKDDEDKRFH-------LFYY----SH 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21 LYLKPLAGVYRSLKKOLENNVMTFNVNVKDILNSRFNKRENFKNVLESDLIPYKDLTSSN 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; DB 1; Length 3135;
0.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 188 SDDTLIEWFDDNTNEENFLLFEKRCL-----MKIFSSPKRKTVVQK 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             228 HNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQ 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SOFTWARE: PATCHLIN Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/323,170B
FILING DATE: 13-0CT-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/010,409
FILING DATE: 29-JAN 1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                     Townsend and Townsend and Crew LLP 5 Embarcadero Center, 8th Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            81 YVVKDPYKFLNKEKRDKFLSS---YNYIKD---SIDTDINF--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               015280-113100US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; Pred. No. 0.98
41; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6.2%; Score 122; 22.4%; Pred. No. 0
                                             1261 MEKLKELEKALSEDSKE----IIDAKDD 1284
197 LNYLKTIQDKLADFKKNNNFVGIADLSTD 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                  Sequence 2, Application US/08323170B Patent No. 5733772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                169 ----IQNVEEKIQRDTYE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: Quine, Jonathan A. REGISTRATION NUMBER: P-41,261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE/DOCKET NUMBER: 01
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 3135 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                              STREET: Two Embarcadero
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; MOLECULE TYPE: protein US-08-323-170B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend
                                                                                                                                                                                                                                                                                                                                                                                                                                     San Francisco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZIP: 94111-3834
                                                                                                                       RESULT 10
US-08-323-170B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                65;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
```

G

```
214 -- NNFVGIADLSTDYNHNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRK 271
272 KTMIQQSSGCFRHLDEREECKCLLNYKQE-----GDKCVENPNPTCNENNGGCDADAK 324
                                                                                                                                                                                             496 RGLV-----LPE----LNYDLEYFNEHLYNDK--NSPEDKDNKGKGVVHVDTT 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62 -FKNVLESDLIPYKDLTSSNYVVKDPYKFLN--KEKRDKFLSSYNYIKDSIDTDINF--- 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     292 PFTSIL---IHAYKEHNGTNLIESKNYALGSDIPEKCDT-LASNCFLSGNFNIEKCFQCA 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           239 LONICETGKNFKLVVYIKENTLIIKWKVYGETKDTTENN----KVDVRKYL---INEKET 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels 103; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7 IDNILSKIENEYEVLYLKPLA-----GVYRSLKKQLENNVMTFNVNVKDILNSRFNKREN 61
                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Paoletti, Enzo
APPLICANT: Tartaglia, James
APPLICANT: Cox, William I
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
NUMBER OF SEQUENCES: 217
                                                                                                      455 IFNNLKRLLIYHSEENINTLKNKFRNAAV------CLKNVDDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/184,009
FILING DATE: 19-JAN-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; Score 121.5; DB 2;
; Pred. No. 0.23;
54; Mismatches 149;
                                                                                                                                                                                                                                        325 CTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSS 358
                                                                                                                                                                                                                                                                                 538 LEKEDTLSYDNSDNMFCNKEYCNRLKDENNCISN 571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Curtis, Morris & Safford
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ., 19-JAN-1994
ON: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                          Sequence 120, Application US/08184009 Patent No. 5833975
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEPHONE: (212) 840-3333
TELEFRAX: (212) 840-0712
TELEX: 475066CURTMS
INFORMATION FOR SEQ ID NO: 120:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 18,100 PELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 6.1%;
Best Local Similarity 22.3%;
Matches 88; Conservative 5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STREET: 530 Fifth Avenue
CITY: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE:
FRAGMENT TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                        US-08-184-009-120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-184-009-120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                             δ
                                                                                                      g
                                                                                                                                                  Ω
                                                                                                                                                                                           g
                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         62 -FKNVLESDLIPYKDLTSSNYVVKDPYKFLN--KEKRDKFLSSYNYIKDSIDTDINF--- 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       348 LLVEKENKNDVC--YKYLSEDIVSKFKEIKAETEDDDEDDYTEYKLTESIDNILVKMFKT 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       --VNDKIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQD-----KLADFKKN---- 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       339 LQNICETGKNFKLVVYIKENTLIIKWKVYGETKDTTENN----KVDVRKYL---INEKET 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----ANDVLGYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIET----LYKT 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7 IDNILSKIENEYEVLYLKPLA----GVYRSLKKQLENNVMTFNVNVKDILNSRFNKREN 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 984;
                                                                                                                                           APPLICANT: Paoletti, Enzo
APPLICANT: de Taisne, Charles
APPLICANT: Tine, John A.
TITLE CON INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
NUMBER OF SEQUENCES: 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757 077 FFITME TATE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54; Mismatches 149;
                                                                                                                                                                                                                                                                             SSEE: Curtis, Morris & Safford, P.C.
F: 530 Fifth Avenue, 25th Floor
New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6.1%; Score 121.5; D 22.3%; Pred. No. 0.23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            454310-2570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/075,783
FILING DATE: 11-UNN-1993
PRIOR APPLICATION NUMBER: US 07/852,305
FILING DATE: 18-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,183
FILING DATE: 20-MAR-1991
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 UMBER: US/08/257,073
09-JUN-1994
                                                                                                                                                                                                                                                                                                                                                                      UNITED STATES OF AMERICA
                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                             Sequence 3, Application US/08257073
Patent No. 5766597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME: Frommer, William S. REGISTRATION NUMBER: 25,506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE/DOCKET NUMBER: 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEPHONE: (312) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066 CURINS
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 22.33
Matches 88; Conservative
                                                                                                                                                                                                                                                           CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: amino acid
STRANDEDNESS: si
                                                                                                                                                                                                                                                                                                                                                    New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; MOLECULE TYPE:
; FRAGMENT TYPE:
US-08-257-073-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: UN
                                                                                                                                                                                                                                                                                 ADDRESSEE:
STREET: 5:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    165
```

20;

셤 ò Q ò qq

```
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miler, Louis H.
APPLICANT: Miler, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Wellems, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
                                                                                                                                                                                                                                      348 LLVEKENKNDVC--YKYLSEDIVSKFKEIKAETEDDDEDDYTEYKLTESIDNILVKMFKT 405
                                                                                                     62 -FKNVLESDLIPYKDLISSNYVVKDPYKFLN--KEKRDKFLSSYNYIKDSIDTDINF--- 115
                                                                                                                                      292 PFTSIL---IHAYKEHNGTNLIESKNYALGSDIPEKCDT-LASNCFLSGNFNIEKCFQCA 347
                                                                                                                                                                                                    116 -----ANDVLGYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIET----LYKT 164
                                                                                                                                                                                                                                                                                                       165 -- VNDKIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQD-----KLADFKKN---- 213
                                                                                                                                                                                                                                                                                                                                                      406 NENNDKSELIKL------EEVDDSLKLELMNYCSLLKDVDTTGTLDNYGMGNEMD 454
                                                                                                                                                                                                                                                                                                                                                                                                       214 -- NNFVGIADLSTDYNHNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRK 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                   455 IFNNLKRLLIYHSEENINTLKNKFRNAAV-----CLKNVDDW-----IVNK 495
                                                 239 LONICETGKNFKLVVYIKENTLIIKWKVYGETKDTTENN----KVDVRKYL---INEKET 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       272 KTMIQQSSGCFRHLDEREECKCLLNYKQE-----GDKCVENPNPTCNENNGGCDADAK 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        496 RGLV------LPE-----LNYDLEYFNEHLYNDK--NSPEDKDNKGKGVVHVDTT 537
     7 IDNILSKIENEYEVLYLKPLA-----GVYRSLKKQLENNVMTFNVNVKDILNSRFNKREN 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA: APPLICATION DATA: US/08/568,459A FILING DATE: 07-DEC-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE: Knobbe Martens Olson 6 Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            325 CTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSS 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     538 LEKEDTLSYDNSDNMFCNKEYCNRLKDENNCISN 571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NIH121.001CP1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 4, Application US/08568459A Patent No. 5849306 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE/DOCKET NUMBER: N
TELECOMMUNICATION INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEO ID NO: 4:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH: 1435 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CORRESPONDENCE ADDRESS ADDRESSEE: Knobbe Ma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-568-459A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STATE:
                                                                                                     δ
                                                                                                                                                   d
                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                       ōλ
                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             20;
-----ANDVLGYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIET----LYKT 164
                                              348 LLVEKENKNDVC--YKYLSEDIVSKFKEIKAETEDDDEDDYTEYKLTESIDNILVKMFKT 405
                                                                                               165 -- VNDKIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQD-----KLADFKKN---- 213
                                                                                                                            --NNFVGIADLSTDYNHNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRK 271
                                                                                                                                                                                                                                                                                                  272 KTMIQQSSGCFRHLDEREECKCLLNYKQE-----GDKCVENPNPTCNENNGGCDADAK 324
                                                                                                                                                                                                                                                                                                                                                   496 RGLV------LPE-----LNYDLEYFNEHLYNDK--NSPEDKDNKGKGVVHVDTT 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 984;
                                                                                                                                                                                                                                 455 IFNNLKRLLIYHSEENINTLKNKFRNAAV------CLKNVDDW-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-458-356-120

Sequence 120, Application US/08458356

Patent No. 5942235;
GAPELICANT: Paoletti, Enzo
APPLICANT: Tartaglia, James
APPLICANT: Cox, William I.
TITLE OF INVERTION: RECOMBINANT VIRUS IMMUNOTHERAPY
NUMBER OF SEQUENCES: 217

CORRESPONDENCE ADDRESS:
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,356
FILING DATE: 02-JUN-1995
CLASSIFICATION: 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 6.1%; Score 121.5; DB 2;
Best Local Similarity 22.3%; Pred. No. 0.23;
Matches 88; Conservative 54; Mismatches 149;
                                                                                                                                                                                                                                                                                                                                                                                                  325 CTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSS 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                   538 LEKEDTLSYDNSDNMFCNKEYCNRLKDENNCISN 571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PULDADILITATION:
PULDADILITATION DATA:
APPLICATION NUMBER: US 08/184,009
FILING DATE: 19-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFRERENCE/POCKET NUMBER: 444310-2530
TELEPHONE: (212) 840-3333
TELEPHONE: (212) 840-3333
TELEPA: 42506CURTMS
INFORMATION FOR SEQ ID NO: 120:
SEQUENCE CHARACTERISTICS:
LENGTH: 984 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             peptide
internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRANDEDNESS: single
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; MOLECULE TYPE:
; FRAGMENT TYPE:
US-08-458-356-120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                           셤
                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                  ò
```

```
TELECOMMUNICATION INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            325
                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97
                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Su, Xin-zhaun
APPLICANT: Wellems, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
                                                                                                                                                               19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   220 LKKNDN------KYNSKFCNDLKNSFLDYGHLAMGNDMDFGGYSTKAENKIQEVFKG 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     300
                                                                                                                                                                                                                                                                                                                79 WKSYGTPDNIDKNMSLINKHNNEEMFNNNYQS-----FLSTSSLIKQNKYVPINAVRVS 132
                                                                                                                                                                                                                                                                                                                                                                                                                                            154 FLNNIETLYKTVND-KIDLFVIHLEAKVLNYTYEKSN---VEVKIKELNYLKTIQDKLAD 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----FENLLKS 248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          271 AHGEISEHKIKNFRKEWWNEFREKLWEAMLS---EHKNNINNCK---NIPQEELQITQWI 324
                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                   -----ENFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRD 96
                                                                                                                                                                                                      -- ENNVMTFNVNVKDIL-NSRFNKR 59
                                                                                                                                                                                                                                           19 AKARNEYDIKENEKFLDVYKEKFNELDKKKYGNVOKTDKKIFTFIENKLDILNNSKFNKR 78
                                                                                                                                                                                                                                                                                                                                                               97 KFLSSYNYIKDSIDTDINFANDVLGYYKILS---EKYKSDLDSIKKYINDKQGENEKYLP
                                                                                                                                                                                                                                                                                                                                                                                        249 VLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE----
                                                                                                                     Query Match 6.0%; Score 119.5; DB 2; Length 1435; Best Local Similarity 21.0%; Pred. No. 0.55; Matches 81; Conservative 46; Mismatches 123; Indels 135;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----YVCIPDRRIQLCIVNLSI-IKTYTKETMKDHFIEASKKESQLL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE: Knobbe Martens Olson & Bear STREET: 620 Newport Center Drive 16th Floor CITY: Newport Beach STATE: California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FKKNNNFVGIADLSTDYNH - - - NNLLTKFLSTGMV - - - -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
HYPOTHETICAL: NO
ORGANISM: Plasmodium falciparum
US-08-568-459A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----GDKCVEN-----PNPTCNEN 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       325 KEWHGEFLLERDNRSKLPKSKCKNN 349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 4, Application US/08487826B Patent No. 5993827
GENERAL INFORMATION:
                                                                                                                                                                                                    12 SKIENEYEVLYLKPLAGVYRSLKKQL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Peterson, David S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10-SEP-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      92660
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-487-826B-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                      음
                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ç
                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                 ò
```

```
19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --FENLLKS 248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            178 -----YVCIPDRRIQLCIVNLSI-IKTYTKETMKDHFIEASKKESQLL----- 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20 LKKNDN-----KYNSKFCNDLKNSFLDYGHLAMGNDMDFGGYSTKAENKIQEVFKG 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       79 WKSYGTPDNIDKNMSLINKHNNEEMFNNNYQS-----FLSTSSLIKQNKYVPINAVRVS 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KFLSSYNYIKDSIDTDINFANDVLGYYKILS---EKYKSDLDSIKKYINDKQGENEKYLP 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       154 FLINIETLYKTVND-KIDLFVIHLEAKVLNYTYEKSN---VEVKIKELNYLKTIQDKLAD 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          249 VLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE----- 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             271 AHGEISEHKIKNFRKEWWNEFREKLWEAMLS---EHKNNINNCK---NIPQEELQITQWI 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12 SKIENEYEVLYLKPLAGVYRSLKKQL-----ENNVMTFNVNVKDIL-NSRFNKR 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19 AKARNEYDIKENEKFLDVYKEKFNELDKKKYGNVQKTDKKIFTFIENKLDILNNSKFNKR 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -- ENFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRD 96
                                                                                                                                                                                                                                                                                                                                                                                                                       Length 1435;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    133 RILS-----FLDSRINNGRNTSSNNEVLSNCREKRKGMKWDCKKK-NDRSN---
                                                                                                                                                                                                                                                                                                                                                                                                                       DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                 6.0%; Score 119.5; DB 2;
21.0%; Pred. No. 0.55;
tive 46; Mismatches 123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    210 FKKNNNFVGIADLSTDYNH---NNLLTKFLSTGMV-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8, 2001, 12:30:04
                                                                                                                                                                                                                                                                                                        Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KEWHGEFLLERDNRSKLPKSKCKNN
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEO ID NO: 4:
SEQUENCE CHARACTERISTICS:
                                                                                                            LENGIH: 1435 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                        TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
                                                                                                                                                                                                                     MOLECULE TYPE: protein HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Search completed: August
Job time: 52 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         301 ----GDKCVEN---
                                                                                                                                                                                                                                               HYPOTHETICAL: I
                                                                                                                                                                                                                                                                                                        ), ORGANISM:
US-08-487-826B-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        81;
                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
```

Ē

plasmodium

plasmodium plasmodium plasmodium plasmodium plasmodium plasmodium plasmodium plasmodium plasmodium plasmodium

plasmodium plasmodium plasmodium

95

plasmodi plasmodi

plasmodium plasmodium plasmodium

025970 025980 025980 025988 025578 025572 025572 025725 025727 025772 025773 025773 025773 025773 025773 025773 025773 025773 025773 025773

plasmodium plasmodium

```
MEDLINE-93295445; PubMed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
Jongwutiwes S., Tanabe K., Kanbara H.;
Jongwutiwes S., Tanabe K., Kanbara H.;
Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates.";
Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13359; BAA02620.1;
InterPro; IPR000561;
InterPro; IPR001245;
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 89.9%; Score 1767.5; DB 5; Length 539; Best Local Similarity 91.8%; Pred. No. 6.8e-89; Matches 346; Conservative 7; Mismatches 21; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               539 AA; 61017 MW; 5CA9C651BB62B5E1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TREMBLrel. 01, Last sequence update)
01-UNN-2000 (TrEMBLrel. 14, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                539 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ALIGNMENTS
025967
025970
025982
025988
025988
025988
025578
025572
025725
043997
043997
025773
043997
025773
043997
025773
043997
043997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q25718
Q25728
Q9UB87
Q9NCN2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
   $\circ$ \circ$ \
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=5833;
1097.5
1091.5
1091.5
1091.5
1090.5
1009
1076
1012.5
1009.5
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
1009.3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  025971;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        025971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
025971
       SET THE PROPERTY OF THE PROPER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                     (without alignments)
708.069 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 025971 plasmodium 025973 plasmodium 02596 plasmodium 025976 plasmodium 025976 plasmodium 025972 plasmodium 025924 plasmodium 025974 plasmodium 025977 plasmodium 025977 plasmodium 025979 plasmodium 025989 plasmodium 025989 plasmodium 025989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AVTPSVIHNILSKIENEYEV.....SNFLGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                          ; Search time 70.07 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
                                      GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    425026 segs, 132305027 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                                          August 8, 2001, 12:33:44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            025571
025573
025966
025596
025597
025597
025972
025973
025977
025977
025977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_rodent:*
sp_unclassified:*
sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_human:*
sp_invertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Post-processing: Minimum Match 08
                                                                                                                                                                                                                                                                                                                                                                                                 US-09-500-376-4
1967
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_mammal:*
sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_plant:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_fungi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SPTREMBL 16:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            569
1694
1694
1704
569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        \begin{array}{c} 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1767.5
1767.5
1763.5
1763.5
1763.5
1763.5
1760.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Scoring table:
                                                                                                                                                                                     protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database
                                                                                                                                                                                                                                                          Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Result
```

5

э; т

~

```
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25
25
287
383
400 /
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Malaria;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NON_TER
SIGNAL
                                                                                                                                                                                                                                                003999
                                                                                                                 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   299
                                                                                                                                                                                                             RESULT
Q03999
                                       δ
                                                                          g
                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ä
                                                                                             GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCSSSNF 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VFENILKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Plasmodium falciparum from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE 60
                   SEQUENCE FROM N.A.
MEDLINE-93295445; PubMed-8515786;
MEDLINE-93295445; Tanaboe K., Kanbara H.;
Noquence conservation in the C-terminal part of the precursor to major merozoite surface proteins (MSP1) of Plasmodium falciparum field isolates.";
GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV
                                                                            LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                    VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                                                  3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01, Last sequence update)
14, Last annotation update)
PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 1767.5; DB 5;
Pred. No. 6.8e-89;
7; Mismatches 21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           105C2E02974FB398
                                                                                                                                                                                                                                                                                                                                                                                                                                      539 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mol. Biochem. Parasitol. 59:95-100(1993)
EMBL; D13361; BAA02622.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROSITE; PS00109; PROTEIN_KINASE_TYR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .... = ==::- ==
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61047 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  89.9%;
91.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                     LGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Merozoite; EGF-like domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MAJOR MEROZOITE SURFACE MSP1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                          LGISFLLILMLILYSFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR000561; -
InterPro; IPR001245; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00008; EGF; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
Matches 346; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         539 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1996
01-JUN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                  Q25973
Q25973;
                                                                                                                                                    239
                                                                                                                                                                                                                                                                                                     359
 121
                                                                          181
                                                                                                                                                                                                                            299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               239
                                                                                                                                                                                                                                                                                                                                                                                                                SOTATE THE SOUND S
                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                      g
 ð
                               g
                                                                          õ
                                                                                                       g
                                                                                                                                                ò
                                                                                                                                                                                                                          οχ
                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
299 GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCSSSNF 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          85 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 323
404 VFENLAKTVLSNLLDGNL-OGMLNISQHQCVKKQCPQNSGCFRHLDEREECKCLLNYKQE 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BY A GPI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VFENILIKSILSNILDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCSSSNF 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25 AVTPSVIDNILSKIENEXEVLYLKPLAGVYRSLKKQLENNVMTFNVVKDILNSRFNKRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -!- MISCELLANEOUS: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE 83K, 42K & 19K ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS C MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           463 GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181 LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam, PF00008; EGF; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 1763.5; DB 5; Length 400;
Pred. No. 8.1e-89;
7; Mismatches 22; Indels 3:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Blackman M.J., Ling I.T., Nicholls S.C., Holder A.A.;
Submitted (XXX-1991) to the EMBL/GenBank/DDBJ databases.
-1- SUBCELLULAR LOCATION: COULD BE ATTACHED TO THE MEMBRANE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           83K MEROZOITE SURFACE ANTIGEN.
42K MEROZOITE SURFACE ANTIGEN.
19K MEROZOITE SURFACE ANTIGEN.
                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
MEROZOITE SURFACE ANTIGEN PRECURSOR 1 (PMMSA) (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEMBRANE ANCHOR.
537F075058626AC2 CRC64;
                                                                                                                                                                                                                                                                                                                                               Š
                                                                                                                                                                                                                                                                                                                                         400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL.
                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MΣ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               89.7%;
91.5%;
                                                                                                                                                        LGISFLLILMLILYSFI 375
                                                                                                                                                                                   523 LGISFLLILMLILYSFI 539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; M64681; AAA29709.1;
InterPro; IPR000561; -.
InterPro; IPR001245; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      45824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GPI-anchor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 91.5
Matches 345; Conservative
                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   24
400
286
400
400
```

셤 ð 

```
Jongwutiwes S., Tanabe K., Kanbara H.; "Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   164 AVTPSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNKDILNSRFNKRE 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LNYTYEKSNVE--IKELLYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8BF9C64322E9A778 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-UIN-2000 (TrEMBLrel. 14, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMEN)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 89.7%; Score 1763.5; DB Best Local Similarity 91.5%; Pred. No. 1.1e-88; Matches 345; Conservative 7; Mismatches 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     539
                                                                                                                                                                                                                                                                                                                                                                                                                                     Mol. Biochem. Parasitol. 59:95-100(1993)
EMBL; D13362; BAA02623.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF00008; EGF; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
                                                       01, Created)
01, Last seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                            MEDLINE-93295445; PubMed-8515786;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           539 AA; 61045 MW;
     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Merozoite, EGF-like domain.
NON_TER 1 1
SEQUENCE 539 AA; 61045 M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR000561; -.
                                                                                                                                                                                   Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001245;
                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=5833;
                                                  01-NOV-1996 (
01-NOV-1996 (
01-JUN-2000 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        025984;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   025984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
     ID DIA POLICE ON SOCIO ON SOCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-93295445; PubMed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
Sequence conservation in the C-terminal part of the precursor to major merozoite surface proteins (MSPI) of Plasmodium falciparum field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ж
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7B7DE90C1D0ACDE7 CRC64;
                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, Created)
1-NOV-1996 (TREMBLrel. 01, Last sequence update)
01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 1763.5; DB
Pred. No. 1.1e-88;
7; Mismatches 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mol. Biochem. Parasitol. 59:95-100(1993)
EMBL; D13357; BAA02618.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF00008; EGF; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61144 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 91.5%;
Matches 345; Conservative
                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Merozoite; EGF-like domain.
                                                                                                    384 LGISFLLILMLILYSFI 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR000561; -. InterPro; IPR001245; -.
                                                                            LGISFLLILMLILYSFI
                                                                                                                                                                                                                                                                                                                                                                                                                                   Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             539 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
SEQUENCE
                                                                                                                                                                                                                                                       025966
                     324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61
```

5

3; Gaps

Indels

22;

5; Length 539;

462

121 284 181

셤

g

ò g ò 239 404 299 463 359

g

ò

à

Q δ g

à

RESULT Q25976

AA.

```
Pfam; PF00008; EGF; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
IPR001245;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota;
                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROSITE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             025972
                                                                                                                                                                                                                                                                                                                                                                                                                                         æ
                                                                                                                                                                                                                                                                                                                                             299
                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
025972
 DR
DR
FT
SO
                                                                                                                             ŏ
                                                                                                                                                                     ö
                                                                                                                                                                                          g
                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;;
                 MEDINE-93295445; PubMed-8515786;
Jongwutiwes S., Tanabe R., Kanbara H.;
"Sequence conservation in the C-terminal part of the precursor to the
                                                 major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates.";
                                                                                                                                                                                                                                               463 GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSNF 522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDINE-93295445; Pubmed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal part of the precursor to the
                                                                                                                                                                                                                                                                             NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                       121 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                 LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                   299 GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCSSSNF 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Plasmodium falciparum from
                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                    1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE 60
                                                                                                                                                                                                                                                                                                                                  239 VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                'n
                                                                                                                                                                                           539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                          Length
                                                                                                                                                                                                                22; Indels
                                                                                                                                                           61114 MW; 3788015F3127CB9E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13, Created)
13, Last sequence update)
14, Last annotation update)
PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                           DB 5;
                                                                                                                                                                                          Score 1763.5; DB Pred. No. 1.1e-88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     539 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               major merozoite surface proteins (MSP1) of
field isolates.";
Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL: D13358; BAA02619.1;
                                                                                                                                                                                                               7; Mismatches
                                                                     Mol. Biochem. Parasitol. 59:95-100(1993)
EMBL, D13356, BAA03617.1; -.
InterPro; IPR000561; -.
InterPro; IPR001245; -.
                                                                                                                 Pfam; PF00008; EGF; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR;
                                                                                                                                                                                          89.7%;
91.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (TrEMBLrel. 13, (TrEMBLrel. 13,
                                                                                                                                       Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        359 LGISFLLILMLILYSFI 375
                                                                                                                                                                                                    Best Local Similarity 91.5
Matches 345; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-MAY-2000 (TrEMBLrel.
01-MAY-2000 (TrEMBLrel.
01-JUN-2000 (TrEMBLrel.
MAJOR MEROZOITE SURFACE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Plasmodium falciparum
                                                                                                                                                           539 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=5833;
                                                                                                                                                 NON_TER
SEQUENCE
                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9TYG1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 523
                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               <sub>.</sub> පු
염
                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                 οŽ
                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
Jongwutiwes S., Tanabe K., Kanbara H.; "Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of Plasmodium falciparum from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             404 VFENLAKTVLSNLLDGNL-OGMLNISQHQCVKKQCPQNSGCFRHLDEREECKCLLNYKQE 462
                                                                                                                                                                                                                                                                                                                                                             61 NFKNVLESDLIPYKDLISSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     284 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFINNIETLYKTVNDKIDLFVIHLEAKV 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCSSSNF 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   463 GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSNF 522
                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                      344 LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 403
                                                                                                                                                                                                                                                    1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE 60
                                                                                                                                                                                                                                                                                                                                                                                            224 NFKNVLESDLIPYKDLISSNYVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL
                                                                                                                                                                                                 3;
                                                                                                                                              Length 539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5; Length 539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                           C70C2E100EC4A101 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              72186F1412C57D1C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                              5;
                                                                                                                                           89.7%; Score 1763.5; DB 91.5%; Pred. No. 1.1e-88; Live 7; Mismatches 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 1760.5; DB Pred. No. 1.6e-88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           539 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             field isolates.";
Mol. Biochem. Parasitol. 59:95-100(1993)
EMBL; D13360; BAA02621.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROTEIN_KINASE_TYR; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE-93295445; PubMed=8515786;
                                                           61075 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61076 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              89.5%;
91.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   359 LGISFLLILMLILYSFI 375
                                                                                                                                                                                                 Matches 345; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-NOV-1996 (TrEMBLrel. 01-NOV-1996 (TrEMBLrel. 01-JUN-2000 (TrEMBLrel. 301-JUN-2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR000561; -. InterPro; IPR001245; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               523 LGISFLLILMLILYSFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF00008; EGF; 1.
                                                           539 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              539 AA;
                                                                                                                                                                          Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PS00109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=5833;
```

```
GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                             Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 651 AA; 74134 MW; AA2137E699255150 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
MEROZOITE SURFACE ANTIGEN 1 (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          56.6%; Score 1113.5; DB 5;
57.0%; Pred. No. 2.9e-53;
iive 59; Mismatches 85; I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       651 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               merozoite surface antigen-1.";
Exp. Parasitol. 81:47-54(1995).
EMBL; 235329; CAA84558.1; -.
InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
Merozoite.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (TrEMBLrel. 01,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1996 (TrEMBLrel. 01,
                                                                                                                                                                                                                                                                                                                                                                                                                     LGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                             [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN=RO-71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN=RO-71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Simi
Matches 221;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               025924;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       025924
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10
                                                   284
                                                                                                      181
                                                                                                                                                         344
                                                                                                                                                                                                          239
                                                                                                                                                                                                                                                           404
                                                                                                                                                                                                                                                                                                                                                                                                                     359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 386
  121
                                                                                                                                                                                                                                                                                                              299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACCOOK NAME OF THE STATE OF THE
                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                        Q
                                                                                                                                                    g
                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5;
  ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                               GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                      238
                                                                                                                                                                                                                                                                                                                                                                                            462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VFENLLKSILSNILDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE 60
  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ţ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCSSSNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE=93295445; PubMed=8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
Sequence conservation in the C-terminal part of the precursor to major merozoite surface proteins (MSP1) of Plasmodium falciparum field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3;
  3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22; Indels
22; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3EED87473EE87B65 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
MSP1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 1760.5; DB 5;
Pred. No. 1.6e-88;
8; Mismatches 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               539 AA
  Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mol. Biochem. Parasitol. 59:95-100(1993)
EMBL; D13363; BAA02624.1; -.
InterPro; IPR001561; -.
InterPro; IPR001245; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PROTEIN_KINASE_TYR; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
  .;
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61046 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 89.5%;
Best Local Similarity 91.2%;
Matches 344; Conservative 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LGISFLLILMLILYSFI 539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LGISFLLILMLILYSFI 375
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00008; EGF; 1
PROSITE; PS00109; PRO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     539 AA;
Matches 344;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q25981
Q25981;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             463
                                                                                                   164
                                                                                                                                                    61
                                                                                                                                                                                                       224
                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                           284
                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                               344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SO DE LA COMPANIA DEL COMPANIA DE LA COMPANIA DEL COMPANIA DE LA COMPANIA DEL COMPANIA DE LA COMPANIA DEL CO
                                                                                                 셤
                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                   ð
                                                                                                                                                    ò
                                                                                                                                                                                                                                                        ά
```

```
298
                                                                                                                                                                462
                                                                                                                                                                                                                        GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCSSSNF 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KYKSDLDSIKKYI----------NDKQGENEKYLPFLNNIETLYKTVNDKI 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9 NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=95354793, PubMed=7628566;
Tolle R., Bujard H., Cooper J.A.;
"Plasmodium falciparum: variations within the C-terminal region of
                             463. GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSNF
LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                            VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=92275047; PubMed=1592091;
Olafsson P., Matile H., Certa U.;
"Plasmodium falatiparum: the repetitive MSA-1 surface protein of
RO-71 isolate is recognized by mouse antibody against the
nonrepetitive repeat block of RO-33.";
Exp. Parasitol. 74:381-389(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
```

4.,

```
Q25974
Q25974;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q25975
Q25975;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    482
                                                                                                                                                                                                                                                                                                                                        6
                                                                                                                                                                                                                                                                                                                                                                          69
                                                                                                                                                                                                                                                                                                                                                                                            244
                                                                                                                                                                                                                                                                                                                                                                                                           129
                                                                                                                                                                                                                                                                                                                                                                                                                              304
                                                                                                                                                                                                                                                                                                                                                                                                                                                               364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   288
                                                                                                                                                                                                                                                                                                                                                                                                                                               170
                                                  RESULT
Q25974
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          025975
                                                                              q
                                                                                                                                                                                                                                                                                                                                                        g
, Ω
                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                               ă
                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DATA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                           287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KYKSDLDSIKKYI-------NDKQGENEKYLPFLNNIETLYKTVNDKI 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPL 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            541
     DLFVIHLEAKVLNYTYEK--SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHN 227
                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
               DMLGKLLSTGLV-QNFPNTIISKLIEGKF-QDMLNISQHQCVKKQCPQNSGCFRHLDERE
                                                                                                                                                                                                                                                                                                      ಭ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DLFVIHLEAKVLNYTYEK -- SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NLLTKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDERE
                                                                                                                                                                                                                                                                                                                                                                                                                                              NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NLLTKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDERE
                                                                                                                                                                                                                                                                                  MEDLINE-93295445; PubMed-8515786;
Angoquatiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal part of the precursor to major merozoite surface proteins (MSPI) of Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                           Length 569;
                                                                                                                                                                                                                                      Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                             85; Indels
                                                                                                                                                                                                                                                                                                                                                                                 A26C45914AF01396 CRC64;
                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
MSP1.
                                                                                                                                                                                                                                                                                                                                                                                                           5
                                                                                                                                                                                                                                                                                                                                                                                                          Score 1110.5; DB :
Pred. No. 3.7e-53;
0; Mismatches 85.
                                                                                                                                                                            569 AA.
                                                                                                                                                                                                                                                                                                                    field isolates.";
Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13346; BAA02607.1; -.
InterPro; IPR000561; -.
Pfam; PF000008; EGF.1.
Merozoite; EGF-like domain.
                                                                                                                        624 FDGIFCSSSNFLGISFLLILMLILYSFI 651
                                                                                                               SMVIFCSSSNFLGISFLLILMLILYSFI 375
                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                            60;
                                                                                                                                                                                                                                                                                                                                                                                  64485 MW;
                                                                                                                                                                                                                                                                                                                                                                                                          56.5%;
56.7%;
                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                 569 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                         NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Sim
Matches 220;
                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                          NON_TER
                                                                                                                                                                           025969
                                                                                                              348
                       446
                                        228
                                                          909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           482
                                                                                           g
                                                                                                                                                                                    ద
                                       ç
                                                         g
                                                                            ò
                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
```

```
Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----NDKQGENEKYLPFLNNIETLYKTVNDKI 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DLFVIHLEAKVLNYTYEK--SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHN 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPL 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    23; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NLLTKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDERE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 1110.5; DB 5; Length
Pred. No. 3.7e-53;
0; Mismatches 85; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   64457 MW; A2691FE569E80396 CRC64;
                                                                                                                                                                      01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Æ
                                                                                                                        569 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    269
                                                                                                                                                                                                                                                                                                                                                                                                            Mol. Biochem. Parasitol. 59:95-100(1993)
EMBL; D13348; BAA02609.1; -.
   375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               348 SMVIFCSSSNFLGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   542 FDGIFCSSSNFLGISFLLILMLILYSFI 569
                                                                                                                                                    01, Created)
01, Last seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                        PRT;
PubMed=8515786;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    60;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   56.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01,
01,
13,
                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (TrEMBLrel.
                                                                                                                                                      (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (TrEMBLrel.
                                                                                                                                                                                                                                           Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KYKSDLDSIKKYI---
                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR000561;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AA;
                                                                                                                                                                                                                                                                                             [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                              MEDLINE=93295445;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1996 (
01-NOV-1996 (
01-MAY-2000 (
                                                                                                                                                        01-NOV-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches 220;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON_TER
SEQUENCE
```

us-09-500-376-4.rspt

```
569 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AA;
                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          569
                                                               NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              025979:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 025979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 15
                                                                                                                                                                                                                                                                                                                                 228
                                                                                                                                                                                                                                           129
                                                                                                                                                                                                                                                                                                           364
                                                                                                                                                                                                                                                                                                                                                                           288
                                                                                                                                                                                               69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        025979
          RE DR ES
                                                                                                                                                                         g
                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                          Dp
                                                                                                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                     Qγ
                                                                                                                                                                                                                                                                  4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AEQUENCE FROM N.A.
MEDLINE=93295445; PubMed=8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates.";
                                                                                                        art of the precursor to the Plasmodium falciparum from
                                                                                                                                                                                                                                                                                                                                            KYKSDLDSIKKYI--------NDKQGENEKYLPFLNNIETLYKTVNDKI 169
                                                                                                                                                                                                                                                                                                                                                                                       287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             424 DMLGKLLSTGLV-QNPPNTIISKLIEGKF-QDMLNISQHQCVKKQCPENSGCFRHLDERE 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPL 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                     170 DLFVIHLEAKVLNYTYEK -- SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHN 227
                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                    9 NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NLLTKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDERE
                                                                                                                                                                                                                                                                  23;
                                                                                                                                                                                                                                          56.5%; Score 1110.5; DB 5; Length 569; 56.7%; Pred. No. 3.7e-53; ive 60; Mismatches 85; Indels 23;
                                         Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBL_TaxID=5833;
                                                                                                                                                                                                         569 AA; 64459 MW; ED4284B2867C9703 CRC64;
                                                                                         Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal part of the major merozoite surface proteins (MSP1) of Plasmodium field isolates.";
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT).
MSP1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   269
                                                                                                                                        Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13349; BAA02610.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   542 FDGIFCSSSNFLGISFLLILMLILYSFI 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       348 SMVIFCSSSNFLGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                   MEDLINE-93295445; PubMed-8515786;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                             InterPro; IPR000561; -
Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                               Conservative
                             Plasmodium falciparum
                                                                                                                                                                                                                                                    Best Local Similarity
Matches 220; Conserv
                                                                        SEQUENCE FROM N.A.
                                                  NCBL_TaxID=5833;
                                                                                                                                                                                              NON_TER
SEQUENCE
                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             025977;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 14.
Q25977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  025977
                                                                                                                                                                                                                                                                                                                                                                           129 1
                                                                                                                                                                                                                                                                                                                                69
                                                                                                                                                                                                                                                                                                                                                                                                                                          364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            288
         ò
                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                     δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | ::|:|:|| | DDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDDTKK 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                  KYKSDLDSIKKYI---------NDKQGENEKYLPFLNNIETLYKTVNDKI 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       170 DLFVIHLEAKVLNYTYEK--SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHN 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPL 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                          9 NILSKIENEYEVLYLKPLAGVYRSLKKOLENNVMTFNVNVKDILNSPFNKRENFKNVLES 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.

SEQUENCE FROM N.A.

JONGWULIWES S., Tanabe K., Kanbara H.;

Sequence conservation in the C-terminal part of the precursor to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 424 DMLGKLLSTGLV-QNPPNTIISKLIEGKF-QDMLNISQHQCVKKQCPENSGCFRHLDERE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       304 KYKDDLESIKKVIKEEKEFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLVNKI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NLLTKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDERE
                                                                                                                                                                                                                                                                          23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 569;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
NCBI_TaxID=5833;
                                                                                                                                                                                                                          DB 5; Length
                                                                                                                                                                                                                       Score 1110.5; DB 5; Length
Pred. No. 3.7e-53;
); Mismatches 85; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         85; Indels
                                                                                                                                                   A26C45914AF01396 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A2691FE569E80396 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NOY-1000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURRACE PROTEIN PRECURSOR (FRADMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Score 1110.5; DB 5;
; Pred. No. 3.7e-53;
60; Mismatches 85;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ą.
Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13350; BAA02611.1; -.
InterPro: IPR000561; -.
Pfam: PF00008; EGF: 1.
Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13352; BAA02613.1; -.
InterPro; IPR000561; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             542 FDGIFCSSSNFLGISFLLILMLILYSFI 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      348 SMVIFCSSSNFLGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                       56.5%; Sco. 56.7%; Pre-
tive 60;
                                                                                                                                                   64485 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             64457 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         56.5%;
56.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 56.5'
Best Local Similarity 56.7'
Matches 220; Conservative
                                                                                                                                                                                                                                                                          Matches 220; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Plasmodium falciparum
```

```
129 KYKSDLDSIKKYI------NDKQGENEKYLPFLNNIETLYKTVNDKI 169
                                                                                                             424 DMLGKLLSTGLV-QNFPNTIISKLIEGKF-QDMLNISQHQCVKKQCPENSGCFRHLDERE 481
                                               DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                              DLFVIHLEAKVLNYTYEK -- SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHN 227
                                                                                                                                                                                             NLITKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDERE 287
9 NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES 68
                                                                                                                                                                                                                                                                                                        FDGIFCSSSNFLGISFLLILMLILYSFI 569
                                                                                                                                                                                                                                                                                            SMVIFCSSSNFLGISFLLILMLILYSFI 375
                                                                                                                                                                     364
                                                                                                                                                                                                                                                                                           348
                                               69
                                                                                                                                              170
                                                                                                                                                                                             228
                                                                                            ŏ
                                                                                                                  qq
                                                                                                                                              ò
                                                                                                                                                                 q
                                                                                                                                                                                             ò
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                               QQ
                                               ò
```

Search completed: August 8, 2001, 12:33:46
Job time: 274 sec

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

August 8, 2001, 12:32:34; Search time 24.79 Seconds (without alignments) 518.185 Million cell updates/sec. Run on:

US-09-500-376-4 1967 1 AVTPSVIHNILSKIENEYEV......SNFLGISFLLILMLILYSFI 375 Title: Perfect score: Sequence:

Scoring table:

BLOSUM62 Gapop 10.0 , Gapext 0.5

93435 Total number of hits satisfying chosen parameters: 93435 seqs, 34255486 residues Searched:

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries Minimum DB seq length: 0 Maximum DB seq length: 200000000

SwissProt\_39:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description								_	_	Q06366 clostridium	Q04956 plasmodium		-			P75093 mycoplasma					Q49419 mycoplasma	O51578 borrelia bu		P45615 mycoplasma	_	P46081 clostridium	P40957 saccharomyc	P33540 neurospora	Q25490 manduca sex	P36016 saccharomyc	32380	P37609 lactococcus	
SUMMARIES	OI.		MSP1_PLAFW	MSP1_PLAFC	- 1	MSP1_PLAF3	MSP1_PLAFF	MSP1_PLAFM	MSP1_PLAYO	BXEN_CLOBO	BXEN_CLOBU	ATX1_PLAFA	Y704_METJA	IRR1_YEAST		CHEA_BORBU	Y018_MYCPN	YLJ2_CAEEL	SERA_PLAFG	RBP2_PLAVB	DPO1_RICPR	Y328_MYCGE	EX5B_BORBU	TIG_BUCAI	YKDA_MYCCA	Y875_METJA	BXCN_CLOBO	MAD1_YEAST	RPOP_NEUCR	APLP_MANSE	LHS1_YEAST	- 1	- 1	CIN8_YEAST
	DB	П	П	~	7	-	Н	7	7	-	ч	-	-	-	-		П	-		-		-	П	Н	П	-	ч	7	Н		п	-	-	-
	Length		1639	1726	1726	1682	1701	1701	1772	1162	1162	1956	377	1150	971	864	1030	2198	989	1251	867	756	1169	442	655	748	1196	749	896	3305	881	944	922	1038
æ	Ouery		89.7					55.5		7.6	•	•				٠	٠			•	6.5		6.4		٠		٠	٠	6.3	•	٠	6.3	٠	6.2
	Score	63.	1763.5	1110	1109	1102.5	1091	1091	565.5	149.5	146.5		136.5	135	134	133	132	129.5	129	. 129	127.5	127	126	125	125	125	125	124.5	124.5	124	S	123	122	122
	Result No.	-	7	m	4	2	9	7	80	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29	30	31	32	33

P39682 saccharomyc P1222 nicotiana t P02381 saccharomyc Q02455 saccharomyc P36112 saccharomyc P21421 plasmodium P89105 saccharomyc P51819 pharbitis n Q000383 epifagus vi P25805 plasmodium Q09622 caenorhabdi
PR39_YEAST YCF1_TOBAC MRAR_YEAST MLP1_YEAST YKZ6_YEAST YKZ6_YEAST CTR9_PLAFA CTR9_YEAST YCF1_EPIVI CYSP_PLAFA YS83_CAEEL
наннаванана
629 1901 396 1875 540 1024 1077 2009 703 1738 569
000000000000000000000000000000000000000
120.5 120 119.5 119.5 118.5 118.5 118.5 118.5 118.5 118.5 118.5
88888884444444444444444444444444444444

## ALIGNMENTS

Page

```
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121
      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                      5
                                                                                                                                                                                                                                                                                 1255 AVTPSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 1314
                                                                                                                                                                                                                                                                                                                                                             GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSNF 1613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
MEDLINE=86014355; PubMed=2995820;
Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V.,
Nicholls S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T.,
                                                                                                                                                                                                                                                                                                                                                                                                                        GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                          NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Holder A.A.;
Submitted (MAR-1991) to the EMBL/GenBank/DDBJ databases.
-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PTM: MEROZÓITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                             AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-FEB-1996 (Rel. 33, Last sequence update)
01-FEB-1996 (Rel. 33, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VFENILKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GSKCVENPNPICNENNGGCDADAKCTEEDSGŞNGKKITCQCTKPDSYPLSMVIFCSSSNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                  LINKED (GLCNAC. . .) (POTENTIAL).
ADBDECGCOA46322 CRC64;
      (POTENTIAL)
                                                                                                                                                                               Length 1630;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            "Primary structure of the precursor to the three major surface antigens of Plasmodium falciparum merozoites."; Nature 317:270-273(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Plasmodium falciparum (isolate Wellcome).
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                 DB 1;
                                                                                                                                                                             Score 1763.5; DB 1;
Pred. No. 8e-89;
7; Mismatches 22;
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                      MM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (Rel. 05, Created)
                                                                                                                                                                               89.7%;
91.5%;
                                                                                                                      AA; 187289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LGISFLLILMLILYSFI 375
                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
  955
1049
1156
1165
1436
                                                                                                                                                                                                    Similarity
  955
1049
1156
1165
11436
1517
1630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=5848;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (PMMSA) (P195).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Freeman R.R.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13-AUG-1987
01-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MSP1_PLAFW
P04933;
                                                                                                                                                                                                                    345;
                                    CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REVISIONS
  CARBOHYD
                      CARBOHYD
                                                                                                                      SEQUENCE
                                                                                                                                                                               Query Match
                                                                                                                                                                                                    Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 2
MSP1_PLAFW
MSP1_PLAFW
AC P04933
DT 01-FEB
DT 01
                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                 1315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1614
                                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1435
                                                                                                                                                                                                                                                                                                                                          61
  FFFFFFS
                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the between the Swiss Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1324 NFKNVLESDLIPPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 1383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1384 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 1443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      299 GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCSSSNF 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1; Length 1639;
                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL
                                                                                                                                                                                                   InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NKED (GLCNAC. . .) (POT 2C255B6616C87F6E CRC64;
                                                                                                                                                                                                                                                                                                        PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22;
                                                                                                                                                                                                                                                                                                  MEROZOITE SURFACE P
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. 8e-89;
7; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 1763.5;
                                                                                                                                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
                                                                                                                                                            EMBL; X02919; CAA26676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           89.7%;
91.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   187618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1623 LGISFLLILMLILYSFI 1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best_Local Similarity 91.5
Matches 345; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           05,
10,
34,
                                                                                                                                                                            PIR; A24594; A24594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (Rel.
(Rel.
(Rel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1165
1174
1445
1526
1639
                                                                                                                                                                                                                                                               Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13-AUG-1987
01-MAR-1989
01-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 3
MSP1_PLAFC
ID MSP1_PLAFC
AC P04934;
DT 13-AUG-1987
DT 01-MAR-1989
DT 01-OCT-1996
                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
```

ᡢ

```
MSP1_PLAFP
P50495;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                isolate.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Chang
                                                                                                                                                                                                                                                                                                                                                                     MSP1_PLAFP
                                                                                                                                                                                                                   1638
                                                                                                  227
                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                   g
                                                                                                                                                                             ò
                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modifited and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1104-1726 FROM N.A.
MEDLINE-88143999; PubMed-3278296;
Weber J.L., Sim B.K.L., Lyon J.A., Wolff R.;
"Merozolte surface protein sequence from the Camp strain of the human
                                                                                                                                                                                                                                                                                                                                                                                                                               42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KYKSDLDSIKKYI--------NDKQGENEKYLPFLNNIETLYKTVNDK 168
                                                                                                                                                   MEDLINE-86205236; PubMed-3517809; Webber J.L., Leininger W.M., Lyon J.A.; Variation in the gene encoding a major merozoite surface antigen of the human malaria parasite plasmodium falciparum."; Nucleic Acids Res. 14:3311-3323(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES.68
                                                                                                                                                                                                                                                                                                                                                                                                                         PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 4 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                            malaria parasite Plasmodium falciparum.";
Nucleic Acids Res. 16:1206-1206(1988).
-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
  SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Přám. PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 1; Length 1726;
                                                        Plasmodium falciparum (isolate Camp / Malaysia).
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           85; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEROZOITE SURFACE PROTEIN 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DD8AD45FA352BCF3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     56.4%; Score 1110; DB 1;
56.6%; Pred. No. 2.9e-53;
tive 60; Mismatches 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                            Eukaryota; Alveolata; Apicomplexa;
NCBI_TaxID=5835;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MM.
                                                                                                                                 SEQUENCE OF 1-1103 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; X03831; CAA27446.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1613
AA; 196197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ransmembrane; GPI-anchor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PIR; A23386; SAZQGM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL).
                  (PMMSA) (P195)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Simi
Matches 220;
  MEROZOITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              69
SOUR THE TERESTEE THE MAN TO BE COULD CONTROL TO THE SOUR THE SOUR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                          EECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYP 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL).
PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
169 IDLFVIHLEAKVLNYTYEK--SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNH 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        hydropathy profile of the the Uganda-Palo Alto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Exp. Parasitol. 67:1-11(1988).
-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-077-1996 (Rel. 34, Created)
01-077-1996 (Rel. 34, Last sequence update)
01-077-1996 (Rel. 34, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                    NNLLTKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam, PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=57270;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S.P., Kramer K.J., Yamaga K.M., Kato A., Case S.E.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROTEIN 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Plasmodium falciparum (isolate Palo Alto / Uganda)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SURFACE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Plasmodium falciparum: gene structure and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              major merozoite surface antigen (gp195) of
                                                                                                                                                                                                                                                                                                   1698 LFDGIFCSSSNFLGISFLLILMLILYSFI 1726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEROZOITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                           347 LSMVIFCSSSNFLGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-89005525; PubMed-3049134;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; M37213; AAA29611.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GPI-anchor
                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             501
567
638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (PMMSA) (GP195).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Siddiqui W.A.
```

```
non-profit
                                                                                        PIR; S06286; S06286
                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 220; Conserv
                                                                                                                                    ransmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MSP1_PLAFF
P13819;
           modified
                       entities
                                                                                                                                                                    TRANSMEM
                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                           CARBOHYD
                                  or send
                                                                                                                                              SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 6
MSP1_PLAFF
                                                                                                              Pfam;
                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MSP-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            228
 OP DI DE CO
                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  s SWISS-PROT entry is copyright. It is produced through a collaboration ween the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its
                                                                   4;
                                                                                                                                              KDMLGKLLSTGLV-QNFPNTIISKLIEGKF-QDMLNISOHQCVKKQCPENSGCFRHLDER 1637
                                                                                                                                                                                                                                                                                                                            346
                                                                                                                                   DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                   Gaps
                                                                                        68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-88166657; PubMed=3327688;
Certa U., Rotmann D., Matile H., Reber-Liske R.;
"A naturally occurring gene encoding the major surface antigen
precursor p190 of Plasmodium falciparum lacks tripeptide repeats.";
EMBO J. 6:4137-4142(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            merozoite surface antigen-1.";
Exp. Parasitol. 81:47-54(1995).
-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PTM: MEROZÓITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OMEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                   -----NDKQGENEKYLPFLNNIETLYKTVNDK
                                                                                                                                                                                                                          IDLFVIHLEAKVLNYTYEK--SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNH
                                                                                                                                                                                                                                                                                                                EECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYP
                                                                                        NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES
                                                                                                                                                                                                                                                                     NNLLTKFLSTGMVFENLLKSILSNLLDMKLARYVKHFTTPMRKKTMIQQNSGCFRHLDER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SURFACE ANTIGENS)
1613 1613 N-LINKED (GLCNAC. . .) (POTENTIAL.) 1726 AA; 196174 MW; 5B59CEEFA2F9A026 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C-terminal region
                                                                  24;
                                            Length 1726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Plasmodium
                                                                  85; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                  MSP1_PLAF3
P19598; Q25921;
P19598; Q25921;
01-FEB-1991 (Rel. 17, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Plasmodium falciparum (isolate ro-33 / Ghana).
Eukaryota; Alveolata; Apicomplexa; Haemosporida;
NCBL_TaxID=5834;
                                            Score 1109; DB 1;
Pred. No. 3.3e-53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       rolle R., Bujard H., Cooper J.A.;
Plasmodium falciparum: variations within the
                                                                  60; Mismatches
                                                                                                                                                                                                                                                                                                                                                                       LFDGIFCSSSNFLGISFLLILMLILYSFI 1726
                                                                                                                                                                                                                                                                                                                                                          347 LSMVIFCSSSNFLGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-95354793; PubMed-7628566;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 1032-1682 FROM N.A.
                                            56.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 1-1061 FROM N.A.
                                                                  220; Conservative
                                                                                                                                                                              KYKSDLDSIKKYI-----
                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (PMMSA) (P190)
CARBOHYD
          SEQUENCE
                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              between
the Euro
                                                        Best Local
                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 5
MSP1_PLAF3
                                                                                                            1340
                                                                                                                                                        1400
                                                                                                                                                                                                    1460
                                                                                                                                                                                                                                                                    227
                                                                                                                                                                                                                                                                                          1580
                                                                                                                                                                                                                                                                                                               287
                                                                                                                                                                                                                                                                                                                                                                                1698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This
                                                                  Matches
                                                                                        6
                                                                                                                                   69
                                                                                                                                                                              129
                                                                                                                                                                                                                          169
                                                                                                                                                                                                                                                                                                                                     1638
                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                        F.S
                                                                                        ò
                                                                                                            용
                                                                                                                                  ò
                                                                                                                                                   8
                                                                                                                                                                        ò
                                                                                                                                                                                              g
                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                          ò
```

```
In no way
commercial
                    noved. Usage by and for commercial
(See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1595 ECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPL 1654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  69 DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                129 KYKSDLDSIKKYI----------NDKQGENEKYLPFLNNIETLYKTVNDKI 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    23; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-JAN-1990 (Rel. 13, Created)
01-JAN-1990 (Rel. 13, Last sequence update)
01-CT-1996 (Rel. 34, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
(PMMAA).
  ļ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |:| ||:|||| |:|||| | : :||
1417 KYKDDLESIKKVIKEEKEFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLVNKI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               170 DLFVIHLEAKVLNYTYEK--SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9 NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NLLTKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDERE
                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC. .) (POTENTIAL)
(GLCNAC. .) (POTENTIAL)
(GLCNAC. .) (POTENTIAL)
(GLCNAC. .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 1; Length 1682;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
  is
                                                                                                                                                                                                                                                                      Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                              POTENTIAL.
MEMBRANE ANCHOR (POTENTIAL).
N.LINKED (GLCNAC. . .) (POTENTIAL).
as its content
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INKED (GLCNAC. . ) (POT C82A1E159948CAD6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    87;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   56.0%; Score 1102.5; DB 56.7%; Pred. No. 7.2e-53; iive 58; Mismatches 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ą
                      and this statement is not removed. requires a license agreement (See
institutions as long
                                                                         an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1655 FDGIFCSSSNFLGISFLLILMLILYSFI 1682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         348 SMVIFCSSSNFLGISFLLILMLILYSFI 375
                                                                                                                  EMBL; M35727; AAA29715.1; -. EMBL; Y00087; CAA68280.1; -. EMBL; Z35326; CAA84555.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1569
                                                                                                                                                                                                                        InterPro; IPR000561; -
Pfam; PF00008; EGF; 1.
                                                                                                                                                                                                                   IPR000561;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AA;
```

```
287 EECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYP 346
                                                    g
                                                                                                                                                                  g
                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                  42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                         PTM: MEROZOTTE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 4 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                            Peterson M.G., Coppel R.L., McIntyre P., Langford C.J., Woodrow G., Brown G.V., Anders R.F., Kemp D.J.;
"Variation in the precursor to the major merozoite surface antigens of Plasmodium falciparum.":

of Plasmodium falciparum.":
-1. SUBCELLUIAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9 NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIR: A54498; A54498.
Interpro: IrRO000561; -.
Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  55.5%; Score 1091; DB 1; Length 1701; 55.8%; Pred. No. 3.1e-52;
Plasmodium falciparum (isolate FC27 / Papua New Guinea).
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC. . . ) (POTER N-LINKED) (GLCNAC. . ) (POTER N-LIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          87; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROTEIN 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SURFACE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEROZOITE
                                                                                                     SEQUENCE FROM N.A.
MEDLINE=88142999; Pubmed=2449612;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; M19143; AAA29653.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    193719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Transmembrane; GPI-anchor.
SIGNAL 1 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1196
1588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                 NCBI_TaxID=5837;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      991
1089
1196
1588
                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 217;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
```

```
(GLCNAC
                                                                                                               POTENTIAL.
                                                                                                                                                                                                                       N-LINKED
N-LINKED
N-LINKED
             EMBL; X05624; CAA29112.1; -.
                                                                                                Transmembrane; GPI-anchor
                           PIR; A26868; A26868.
PIR; B25120; B25120.
                                                                                                                                                                CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                         CARBOHYD
                                                                                                                                                      CARBOHYD
DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                               KYKSDLDSIKKYI----------NDKQGENEKYLPFLNNIETLYKTVNDK
                                                                                                                                                      IDLFVIHLEAKVLNYTYEK - - SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNH
                                                                                                                                                                                                          NNLLTKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDER
```

ò

g ò

69

ò 셤 227 1555

1495

g

(POTENTIAL)

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
Tanabe K., Mackay M., Goman M., Scaife J.G.;
"Allelic dimorphism in a surface antigen gene of the malaria parasite
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PTM: MEROZÓITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 4 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Polymorphism of the precursor for the major surface antigens of Plasmodium falciparum merozoites: studies at the genetic level."; PHEMD J. 4:3829-3829[1985].
                                                                                                                                                                                    01-AUG-1988 (Rel. 08, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROTEIN 1. . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 1-115 FROM N.A.
MEDILTE-86136024; PubMed=3004972;
MACKAY M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,
Stunnenberg H., Pujard H.;
                                                                                                                                                                                                                                                                                      Guinea).
Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                   Plasmodium falciparum (isolate mad20 / Papua New
Eukaryota; Alveolata; Apicomplexa; Haemosporida;
NCBL_TaxID=70153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEROZOITE SURFACE P.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                         PRT; 1701 AA.
                                                             347 LSMVIFCSSSNFLGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                     PubMed=3079521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REVISIONS TO 1403; 1569 AND 1629.
                                                                                                                                                                                                                                                                                                                                                                                                                    Plasmodium falciparum.";
J. Mol. Biol. 195:273-287(1987).
                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A. MEDLINE-88011243;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL).
                                                                                                                                                                                                                                                     (PMMSA) (P190).
                                                                                                                                                         MSP1_PLAFM
P08569;
                                                                                                                                        MSP1_PLAFM
             1613
```

```
4,
                                                                                                                                                                                                                                                                                                                        STRAIN=17XL;
MEDLINE-88124889; PubMed=2448778;
Burns J.M. Jr., Daly T.M., Vaidya A.B., Long C.A.;
"The 3' portion of the gene for a Plasmodium yoelli merozoite surface
antigen encodes the epitope recognized by a protective monoclonal
antibody.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              168
                                                                                                                                                                                                                                                                                                                                                                                                     DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   the
                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 'KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Proc. Natl. Acad. Sci. U.S.A. 85:602-606(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-JAN-1990 (Rel. 13, Created)
01-JAN-1990 (Rel. 13, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                                                    "Sequence analysis upstream of the gene encoding the precursor to najor merozoite surface antigens of Plasmodium yoelli.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IDLFVIHLEAKVLNYTYEK--SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EECKCLLNYKQEGSKCVENPNPTCNENNGCDADAKCTEEDSGSNGKKITCQCTKPDSYP
                                                                                                                                                                                                                                                                                                  NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KYKSDLDSIKKYI--------NDKQGENEKYLPFLNNIETLYKTVNDK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        227 NNLLTKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDER
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
W; 3FC2EC59AF96EA98 CRC64;
                                                                                                                                                                                                                                             24;
                                                                                                                                                                                ; Score 1091; DB 1; Length 1701;
; Pred. No. 3.1e-52;
61; Mismatches 87; Indels 24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             кикагуоta; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBL_TaxID=5862;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         major merozoite surface antigens of Plasmo
Mol. Biochem. Parasitol. 39:285-288(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LSMVIFCSSSNFLGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-90205979; PubMed-2320061;
                                                                                                           XX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1093-1772 FROM N.A.
                                                                                                                                                                                         55.5%;
                                                                                                      193768
                                                                                                                                                                                Query Match
Best Local Similarity 55.8%
Matches 217; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                        1089
1196
1588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (PMMSA) (230 KDA).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
  991
1089
1196
1588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MSP1_PLAYO
P13828;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Lewis A.P.;
                                                                             CARBOHYD
SEQUENCE
  CARBOHYD
                           CARBOHYD
                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MSP1_PLAYO
                                                                                                                                                                                                                                                                                                                                                                                                                                                          1375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1495
                                                                                                                                                                                                                                                                                                                                                                                                     69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DDT TDD DDT DD
  FT
FT
FT
SO
                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseeisb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1630 KSGLVGEGESKKILAKMLNMDGMDLLGVDPKHVCVDTRD---IPKNAGCFRDDNGTEEWR 1686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1521 LYNTQLAAVKEQIATIEAETNDTNKEEKKKYIPILEDLKGLYETVIGQAEEYSEELQNRL 1580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                33; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9 NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           69 DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KYKSDLDSIKKYI-----ND-KQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STGMVFENLLKSILSNLLDWK----LARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         291 CLLNYKQ-EGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 LNYTYEKSNVEI--KELIYLKTIQDKLADF----KKNNNFVGIADLSTDYNHNNLLTKFL
                                                                                                                                                                                                                                                                                                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 1772;
                                                                                                                                                                                    Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                    -> V (IN REF. 2).
9A6291658EB0F45D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                     28.7%; Score 565.5; DB 1; 34.9%; Pred. No. 1.2e-23; Live 73; Mismatches 144;
                                                                                                                                                                                                                       MEROZOITE SURFACE PR
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                      GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 1162
                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1747 GVFCSSSFMGLSILLIITLIVFN 1770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            350 VIFCSSSNFLGISFLLILMLILYS 373
                                                                                                                                                                                                          EMBL; J03612; AAA29762.1; -. EMBL; J04668; AAA29702.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 34.9%
Matches 134; Conservative
                                                                                                                                                                                                 GPI-anchor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              32,
32,
                                                                                                                                                        PIR; A28121; A28121.
PIR; A45532; A45532.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (Rel.
(Rel.
(Rel.
                                                                                                                                                                                                 Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1995
01-NOV-1995
01-NOV-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BXEN_CLOBO P46082;
                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                              SIGNAL
                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BXEN_CLOBO
 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DA TA TA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οy
```

```
Local Similarity
                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                           NCBI_TaxID=1492;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF01742;
                                                                                                                                                                                                  STRAIN-BL6340
                                                                                                                                                                                                                                                                                                                                TOXICITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Neurotoxin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                           Mashike.
                            BXEN_CLOBU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         112
                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           염
                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                            26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     230 DFLISGGIDYKLLNTNPYWFIDKY-FIDTSKNFEKYKNDYEIKIKNNNYIANSIKLYLEQ 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     71 -----IPYKDLTSSNYVVKDPYKFLNKEKR-DKFLSSY-----NYIKDSIDT---- 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 289 KFKIN-VKDIWELNLSYFSKEFQIMMPERYNNALNHYYRKEFYVIDYFKNYNINGFKNGQ 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          148 NEXYLP-----FLNNIETLYKTVND-----KIDLFVIHLEAKVLNY------T 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ONSGCFRHLDEREECKCLLNYKQEGSKCVENPNP-------TCNENNGGCDA 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14 IENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLESDL--- 70
                                                                                                                             Fujii N., Kimura K., Yokosawa N., Yashiki T., Tsuzuki K., Oguma K., "The complete nucleotide sequence of the gene encoding the nontoxic component of Clostridium botulinum type E progenitor toxin."; J. Gen. Microbiol. 139:79-86(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      573 FKNLGPISLINKKENITIPKIKIDEIPSSMLNFSFKDLSENLFNIYC-KNNF-----Y
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        112 --DINFANDV----LGYYK----ILSEKYKSDL------DSIKKY-IND-KOGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            525 WLKAIFR------NYSLDITETQEISNQFGDTKIIPWIGRALNILNTNNSFVEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----TDYNHNNLLTKFLSTGMVFENLLKSILSNLLDW-KLARYVKHFTTPMRKKTMIQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         468 YLQAQMIDSNDINLSSDFLKVISSKGSLVYSFLNNTMDYLEFIKYDKPIDT---DKKYYK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    320 DAKCTEEDSGSNGKKITCQCTKPDSYPLSMV------IFCSSSNFLGISFLLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            185 YEKS-----NVEIKELIYLKTIQDK-LADFKKN-NNFVGIADLS-----
                                     Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                7.6%; Score 149.5; DB 1; Length 1162;
21.5%; Pred. No. 0.31;
ive 70; Mismatches 159; Indels 155;
                                                                                                                                                                                                                                                                                                                                                                                           96468EDDDAE0F39D CRC64;
 E, NONTOXIC COMPONENT
                                                                                                     STRAIN-MASHIKE;
MEDLINE=93195515; PubMed=8450310;
                                                                                                                                                                                                                                                                                                                                        InterPro; IPR000395;
Pfam; PF01742; Peptidase_M27; 1.
PRINTS; PR00760; BONTOXILYSIN.
                                                                                                                                                                                                                                                                                                                                                                                          1162 AA; 136856 MW;
                                                                                                                                                                                                                                                                                                                            EMBL; D12697; BAA02194.1; -.
 BOTULINUM NEUROTOXIN TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 105; Conservative
                         Clostridium botulinum,
                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | | |:|:
LKKIYYNFL 633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LMLILYSFI 375
                                                 Clostridium.
NCBI_TaxID=1491;
                                                                                                                                                                                               TOXICITY
                                                                                                                                                                                                                                                                                                                                                                              Neurotoxin.
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  625
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             d
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  : | | : : | | : : | 348.IKTLPLSKYNKEIINKPELIVNLINQNNTVLMKSNFYGDGLKGNVDNFYSNYIIPYNLN 407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----IPYKDLTSSNYVVKDPYKFLNKEKR-DKFLSSY-----NYIKDSIDT---- 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              --DINFANDV----LGYYK----ILSEKYKSDL------DSIKKY-IND-KQGE 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         289 KFKIN-VKDIWELNLSYFSKEFQIMMPERYNNALNHYYRKEFYVIDYFKNYNINGFKNGQ 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               274 QQNSGCFRHLDEREECKCLLNYKQEGSKCVENPNP------TCNENNGGCD 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                14 IENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLESDL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DFLISGGIDYKLLNTNPYWFIDKY-FIDTSKNFEKYKNDYEIKIKNNNYIANSIKLYLEQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----LNYTYEKSNVEIKELIYLKTIQDK-LADFKKN-NNFVGIADLS------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----TDYNHNNLLTKFLSTGMVFENLLKSILSNLLDW-KLARYVKHFTTPMRKKTMI
                                                                                                                                                                                                                               Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
Clostridium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Fujii N., Kimura K., Yokosawa N., Oguma K., Yashiki T., Takeshi K., Ohyama T., Isogai E., Isogai H.; "Similarity in nucleotide sequence of the gene encoding nontoxic component of botulinum toxin produced by toxigenic Clostridium butyricum strain BL6340 and Clostridium botulinum type E strain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels 171;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Microbiol. Immunol. 37:395-398(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1162;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1162 AA; 136829 MW; C86E9BE932DA78E4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NEKYLP-----FLNNIETLYKTVND-----KIDLFVIHLEAKV-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              69; Mismatches 151;
BXEN_CLOBU STANDARD; PRT; 1162 AA. 006366; 01-FEB-1995 (Rel. 31, Last sequence update) 01-NOV-1995 (Rel. 32, Last annotation update) BOTULINUM NEUROTOXIN TYPE E, NONTOXIC COMPONENT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7.4%; Score 146.5; Di 21.3%; Pred. No. 0.45;
                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-93360835; PubMed-8355622;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Peptidase_M27; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRINTS; PR00760; BONTOXILYSIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; D12739; BAA02231.1; -. InterPro; IPR000395; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 106; Conservative
                                                                                                                                                                                                       Clostridium butyricum.
```

œ

```
DOMAIN
             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                           7704_METJA
                                                                                                                                                                                                                                                                                                                                  1587
  FT
                                                                                                ΩŽ
                                                                                                                   염
                                                                                                                                                                                                     g
                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                            Ω
                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                       SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2 ATPASES). SUBFAMILY V.
                       -----IFCSSSNF 358
                                        568 ---SFVEEFKNLGPIFLINKKENITIPKIKIDEIPSSMLNFSFKDLSENLFNIYC-KNNF 623
---NYSLDITETQEISNQFGDTKIIPWIGRALNILNTNN---- 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding
DOMAIN 1 35 CYTOPLASMIC (POTENTIAL).
                                                                                                                                                                                                                                                                            Krishna S., Cowan G., Meade J.C., Wells R.A., Stringer J.R.,
Robson K.J.;
"A family of cation ATPase-like molecules from Plasmodium
                                                                                                                                                                                                              Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                                     Cell Biol. 120:385-398(1993).
- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + ORTHOPHOSPHATE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CYTOPLASMIC (POTENTIAL).
PHOSPHORYLATION (PROBABLE)
MAGNESIUM (BY SIMILARITY).
MAGNESIUM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EXTRACELLULAR (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EXTRACELLULAR (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EXTRACELLULAR (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EXTRACELLULAR (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CYTOPLASMIC (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CYTOPLASMIC (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CYTOPLASMIC (POTENTIAL).
                                                                                                                                                         30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
PROBABLE CATION-TRANSPORTING ATPASE 1 (EC 3.6.1.-).
                    319 ADAKCTEEDSG-----SNGKKITCQCTKPDSYPLSMV---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLY-LYS.
POLY-ASN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF00122; E1-E2_ATPase; 4.
PROSITE; PS00154; ATPASE_E1_E2; 1.
                                                                                                                                                                                                                                                               MEDLINE=93132070; PubMed=8421054;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; X65738; CAA46646.1; -.
                                                             359 LGISFLLILMLILYSFI 375
                                                                                  -----YLKKIYYNFL 633
                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           35
58
61
80
407
427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      837
845
863
881
905
                                                                                                                                                                                                       Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         928
952
956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         496
1760
1764
251
251
256
941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR001757;
                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
524 KWLKAIFR-
                                                                                                                                                                                                                                                                                                 "A family o
falciparum.
                                                                                                                                        ATX1_PLAFA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSMEM
DOMAIN
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FRANSMEM
                                                                                                                                                004956;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                  624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              METAL
                                        셤
                                                                                  셤
                                                                                                                                        셤
                   ò
                                                             ò
```

```
14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                       1337 NLIHNLYKKKKYMNNYYDIDEVHLIGNNNNNNNNNNKNNSKEKKLPLKNKMKHIRKNESNDNI 1396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----LSKYKYVHHKNYYYPDSCTNLRKKKNSLFYNLK 1440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1483 TRNFIHSLSEQFAFSNLILSFYIIKNDD---NNVYNKNYIYNKNYIYNKNSICNK--NYI 1537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAINS_JAL-1 / DSM 2661 / ATCC 43067;
STRAINS_JAL-1 / DSM 2661 / ATCC 43067;
STRAINS_JAL-1 / DSM 2661 / ATCC 43067;
MEDLINE=96337999; PubMed=86808087;
Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
Schton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
Schton R.A., Kougherty B.A., Tomb J.F., Addams M.D., Railch C.I.,
Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
Scott J.L., Geoghagen N.S.M., Weidman J.F., Puhrmann J.L., Nguyen D.,
Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
Complete, genome sequence of the methanogenic archaeon, Methanococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  43 TFNVNVKDILNSPFNKRENFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSY 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -KSDLDSIKKYINDKQGENEK 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    151 YLPFLNNIETLYKTVNDKIDLFVIHLEAKVLNYTYEKSNVEIKELIYLK-TIQDKLADFK 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     210 KNNNFVGIADLSTDYNHNNLLTKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRK 269
                                                                                                                                                                                                                                                                  97; Gaps
                                                                                                                                                                                                                                                                                                                                             ---KPLAGVYRSLKKQLENNVM 42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Science 273:1058-1073(1996).
-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
                                                                                                                                                                                            Length 1956;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
                                                                                                                                                                                                                                                                  Indels
1363 1372 POLY-ASN.
1680 1684 POLY-ASN.
1956 AA; 230285 MW; AE708AAE99009335 CRC64;
                                                                                                                                                                                        7.0%; Score 137.5; DB 1;
22.0%; Pred. No. 2.5;
Live 52; Mismatches 107;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             377 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----RHKLKEKKNKKNIKRYK 1603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       103 NYIKDSIDTDINFANDVLGYYKILSEKY----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      270 KTMIQQNSGCFRH-LDEREECKCLLNYK 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                             5 SVIHNILSK---IENEYEV--LYL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; U67517; AAB98705.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HYPOTHETICAL PROTEIN MJ0704
                                                                                                                                                                                Query Match /.u*
Best Local Similarity 22.0%
Matches 72; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Methanococcus jannaschii.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1397 TFNTYTSNNIH---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=2190;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Methanococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7704_METJA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               annaschii
                                                                              SEQUENCE
```

σ

```
HYPOTHETICAL PROTEIN BB0228
                                                                                Query Match b.3%
Best Local Similarity 20.7%
Matches 89; Conservative
EMBL; 246881; CAA86966.1;
SGD; S0001288; IRR1.
DOMAIN 44 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nature 390:580-586(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                      65 7
1150 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1118 IEESSONNSE 1127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   324 TEEDSGSNGK 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           burgdorferi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Y228_BORBU
                                       DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    051246;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Y228_BORBU
                                                                                                                                                                                                                                     103
                                                                                                                                                                                                                                                                                                                                                                                     225
                                                                                                                                                                                                                                                                                                                                                                                                                                    273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STTRS
                                                                                                                                       ò
                                                                                                                                                             셤
                                                                                                                                                                                     δ
                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                    6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lisb-sib.ch).
                                                                                                                                                                                                                         287 IEEKIKTRDINSINKFL-----SETY----NNISN-YEMFNDKTYLEIFVHHLKELLEK 335
                                                                                                                                                                                                            PYKFLNKEKRDKFLSSYNYIKDSIDTD----INFAND------VLGYYKI 125
                                                                                                                                                                                                                                                          LSEKYKS-DLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDK--IDLFVIHLEAKVLN 182
                                                                                                                                                                                     ----- 227
                                                                                                                                       Gaps
                                                                                                                                                           LAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLESDLIPYKDLTSSNYVVKD 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHARACTERIZATION.
MEDLINE-96090137; PubMed-7483852;
Kurlandzka A., Rytka J., Gromadka R., Murawski M.;
"A new essential gene located on Saccharomyces cerevisiae chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ä.Υ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Yeast 11:885-890(1995).
-1- FUNCTION: NOT KNOWN; MAY PARTICIPATE IN CELL-CELL AND/OR CELL-SUBSTRATUM INTERACTIONS.
-1- SUBSELLULAR LOCATION: CYTOPLASMIC.
-1- SUBSELLULAR SOME, TO S. POMBE SPACI7H9.20.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D., Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser, Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D., Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V., Whitehead S.; Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
                                                                                                             DB 1; Length 377;
                                                                                                                                    Indels
                                                             POTENTIAL.
2746A0A96425C285 CRC64;
                                                                                                                                                                         YEKEKNENKLOBEIFNELKDELEKLINKCHNKLRELESQNN 376
                                                                                                                                                                                                                                                                                                          183 YTYEKSNVEIKELIY------LKTIQDKLADFKKNNN 213
                                                                                                                                    60;
                                                                                                                                                                                                                                                                                                                                                                                                                   01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
                                                                                                          Similarity 21.7%; Pred. No. 0.44; Score 136.5; DB Similarity 21.7%; Pred. No. 0.44; St. Mismatches
                                                                                                                                                                                                                                                                                                                                                                                               PRT; 1150 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Saccharomyces cerevisiae (Baker's yeast).
                          POTENTIAL
                                                 POTENTIAL
                                      POTENTIAL
               Transmembrane
                                                                        44630 MW;
                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                         46
87
128
155
              Hypothetical protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
STRAIN-S288C / AB972;
                          26
67
108
135
377 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OR YIL026C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-4932
 TIGR; MJ0704;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          IRR1 PROTEIN
                                                                                                                                    48.
                                                                                                                                                                                                                                                                                                                                                                                              IRR1_YEAST
P40541;
                                                                                                            Query Match
Best Local S
                         TRANSMEM
                                      TRANSMEM
                                                TRANSMEM
                                                             TRANSMEM
                                                                        SECUENCE
                                                                                                                                                                                                                                                                                                                                                                       RESULT 13
IRR1_YEAST
                                                                                                                         Best Loc
Matches
                                                                                                                                                                                                                                                          126
                                                                                                                                                                                                                                                                                                                                 336
                                                                                                                                                           56
                                                                                                                                                                                                           86
                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                              q
 셤
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                          ç
                                                                                                                                                                                                                                                                                                                               셤
```

```
22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | :| : | : | : | : | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | | : | : | | : | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1005 NDLPETGFFENIETEFLLFTVKLKGLMK---LNILDERFASRVALNKEKLGPLFKKIVDD 1061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    949 PADATTWRYSMSEQFPFYLHDNASKVLLKIFLYLESLFAKQVDVQLER-----VADEDANL 1004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                889 LRSVWKPIAAIIGRLNTLVISLAATNETFENINSLFYLKWSACTSLMDIIVAIKIFELKL 948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-----HNNLLTKFLSTGMVFENLLKSILSNLLDWKLARYV---KHFTTPMRKKTM-- 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
STRAIN-ATCC 35210 / B31;
MEDLINE-98065943; PubMed-9403685;
Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
Dougherty B., Tomb J.-F., Fletschman R.D., Richardson D.,
Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
Garland S., Fujil C., Cotton M.D., Horst K., Roberts K., Hatch B.,
Garland S., Venter J.C.;
Genomic sequence of a Lyme disease spirochaete, Borrelia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     54 SPFNKRENFKNVLESDLIPYKDLTSSNYVVKDPYKFLNK------EKRDKFLSSY- 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        790 DPNNSRDD------YKDLTCSLYEL----YINKLTIIGRDYPIEVDEELLQLFL 833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -NYIK-----DSIDTDINFANDVLGYYKILS----EKYKSDLDSIKKYIN----D 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    834 NNFVSRIPIMFQDFDDSTAQEINFKMLVL----LATWNLEKWREIIEKVRDYENSISKD 888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----KIDLFVIHL 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     177 EAKVLNYTY------EKSNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDY 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -IQQN-----SGCFRHLDEREECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKC 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLE----NNVMTFNVNVKDILN--- 53
                                                                                                                                                                                                                                                                                                                                                                                  6.9%; Score 135; DB 1; Length 1150;
20.7%; Pred. No. 1.9;
Live 73; Mismatches 136; Indels 132;
       POLY-GLU.
POLY-ASP.
MW; 89688EA09485AC28 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteria; Spirochaetales; Spirochaetaceae; Borrella
NCBI_TaxID=139;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 144 KQGENEKYLPFLNNIETL------YKTVND-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Borrelia burgdorferi (Lyme disease spirochete)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15-DEC-1998 (Rel. 37, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 971 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
48
70
133008 N
```

```
RRARAR RR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οy
between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                   17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
STRAIN-ATCC 35210 / B31;
MEDLINE-98065943; PubMed-9403685;
Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----PYKFLNKEKRDK----------FLSSY--NYIKDSIDTDI 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            372 LIKGILFGYEFALKEEKGQNFPIALMIKSFKGWLNGLHPIKTLQTSYYINEITNKLEKGI 431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          432 YYFENLIEKYLIFNNHYTLISFIPSHDTEKEMEEBIEKKLMAREIEIKQNPEEFLQFKKD 491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LDSIKKYINDKQGENE-KYLPFLNNIETLYKTVNDKIDLFVIHLEAKVLNYTYEKSNV-- 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -NLLTKFLSTGMVFEN-----LLKSILSNL-------LDWKLARYVK-H 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                               325 ISGINTDLKESI-----FSFGLQNVVE---NKEKEFKNLVFSEL---KNLV-KNKIPKE 371
                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       550 VNLFFKLDFLEKEDYIXLSLFKRALQDLSTKNYSYININNKIQNTLGQINISESYDEDID
                                                                                                                                                                                                                                                                                                                                                                                                                   LAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLESDLIPYKDLTSSNYVVKD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           114 NFANDVLGYYKILSEKY-----KSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----EIKELIYLKTIQDKLADFKKNN------NFVGIADLSTDYNHN--
                                                                                                                                                                                                                                                                                                                 Query Match 6.8%; Score 134; DB 1; Length 971;
Best Local Similarity 19.8%; Pred. No. 1.8;
Matches 74; Conservative 60; Mismatches 102; Indels 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-98438936; PubMed=9765799;
Trueba G.A., Old I.G., Saint-Girons I., Johnson R.C.;
"A cheA cheW operon in Borrelia burgdorferi, the agent of Lyme
                                                                                                                                                                                                                     11 protein.
971 AA; 112959 MW; 088A688D7BBC591A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Borrella burgdorferi (Lyme disease spirochete).
Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
NCBL_TaxID=139;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHEA_BORBU STANDARD; PRT; 864 AA. Q44737; P70857; Q4877; 01-NOV-1997 (Rel. 35, Created) 1. DEC-1998 (Rel. 37, Last sequence update) 30-MAY-2000 (Rel. 39, Last annotation update) CHEMOTAXIS PROTEIN CHEA (EC 2.7.3.-).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Res. Microbiol. 148:191-200(1997).
                                                                                                                                                                      EMBL; AE001133; AAC66621.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             263 FTTPMRKKTMIQQN 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         :| |: : |
670 LLAMLRSKSKLKLN 683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHEA OR BB0669
                                                                                                                                                                                               TIGR; BB0228;
                                                                                                                                                                                                                        Hypothetical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  disease."
                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                   56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        14;
                                                                                                                                                                                                                                                                                                                                                                           J. Bacteriol. 179:552-556(1997).

-I- FUNCTION: INVOLVED IN THE TRANSMISSION OF SENSORY SIGNALS FROM THE CHEMORECEPTORS TO THE FLACELLAR MOTORS. CHEA IS AUTOPHOSPHORYLATED; IT CAN TRANSFER ITS PHOSPHATE GROUP TO EITHER CHEB OR CHEY (BY SIMILARITY).

-I- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).

-I- SUBLIARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KDSIDTDINFANDVLGYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTV 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            166 NDKIDL--FVIHLEAKVLNYTYEKSNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTD 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 SVIHNILSKIENEYEVLYLK -----PLAGVYRSLKKQLENNVMTFNVNVKDILNSP 55
       Σ
                                                                                 В.,
Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson I van Vugt R., Pelmer N., Adams M.D., Gocayne J.D., Weidman J., Utterback T., Warthey L., McDonald L., Artiach P., Bowman C., Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B., Smith H.O., Venter J.C.;
                                                                                                                                                                                                                                                                                                                                    "An unexpected flaA homolog is present and expressed in Borrelia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 1; Length 864;
8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         transduction; Transferase; Kinase; Phosphorylation;
                                                                                                                                  "Genomic sequence of a Lyme disease spirochaete, Borrelia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FNKRENFKNVLESDLIPYKDL-TSSNYVVKDPYKFLNKEKRDKFLSSYN---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          L -> I (IN REF. 1).
S -> H (IN REF. 1).
S -> G (IN REF. 1).
L -> S (IN REF. 1).
L -> S (IN REF. 1).
A -> S (IN REF. 1).
S -> P (IN REF. 1).
3; C1111DE0877BE624 C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. 1.8; ; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6.8%; Score 133;
22.1%; Pred. No. 1.
Live 53; Mismatche
                                                                                                                                                                                                                                                                                     MEDLINE=97144545; PubMed=8990312;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; U28962; AAB96835.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AE001168; AAC67024.1; -
EMBL; U62900; AAC44771.1; -.
EMBL; X91907; CAA63002.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRINTS; PR00344; BCTRLSENSOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                51
11
182
187
239
362
559
98352 M
                                                                                                                                                                                                                                   SEQUENCE OF 1-30 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR000410; -.
InterPro; IPR002545; -.
InterPro; IPR002570; -.
Pfam; PF01584; CheW; 1.
Pfam; PF01627; Hpt; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity 22.1 nes 73; Conservative
                                                                                                                                                                               Nature 390:580-586(1997).
                                                                                                                                                                                                                                                                                                               Ge Y., Charon N.W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       rigk; BB0669;
                                                                                                                                                             burgdorferi.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KINASES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Chemotaxis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MOD_RES
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sensory
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                106
```

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

August 8, 2001, 12:30:54 ; Search time 43.63 Seconds .. (without alignments) 654.721 Million cell updates/sec Run on:

Title: Perfect score: Sequence:

US-09-500-376-4 1967 1 AVTPSVIHNILSKIENEYEV......SNFLGISFLLIIMLILXSFI 375

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

219241 seqs, 76174552 residues

Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

PIR\_68:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	major merozoite su	probable major sur	w	major merozoite su	merozoite surface	major merozoite su		major merozoite su	major merozoite su	020	major blood-stage		major merozoite su	major merozoite su	, ORF MSV152 probabl	rhoptry protein -	ORF MSV156 hypothe	Ser/Thr protein ki	hypothetical prote	progenitor toxin n	dihydropteroate sy	hypothetical prote	botulinum toxin no	VPS45-like protein	rhoptry protein -	DNA-directed RNA p	hypothetical prote	ORF MSV140 hypothe	phosphatase (acid
SUMMARIES	QI	SAZOK1	A24594	A45545	805603	S47282	SAZOGM	A45948	A54498	A26868	A39401	A45604	A28121	A45532	A45546	28313	T28677	T28317	E71609	671613	A47708	A57719	A71622	140817	D71607	28676	S72284	E71606	œ	71616
	DB I	1 S	2 A		2 S		1 S	2 A	2 A			2 A	2 A	2 A	2 A			2 T		_	-								2 T	2 B
	Query Match Length D		1640	400	1639	651	1726	1726	1701	1701	1726	1751	680	1772	1785	1306	2269	1127	1714	1346	1162	206	202	1162	722	2401	096	980	809	2010
æ	Query	97.3	96.6	89.7	89.7	9.95	56.4	56.1	55.5	54.8	36.3	36.2		28.7	26.9	8.4	8.4	8.0	7.9	7.8	7.7	7.6	7.5	7.4	7.4	7.4	7.4	7.4	7.3	7.2
	Score	1914	1901	1763.5	1763.5	1113.5	1110	1103	1001	1077	714	713	565.5	565.5	529	165.5	164.5	157	156	153	150.5	149.5	7	146.5	14	145.5	145	145	14	142.5
	Result No.	-	7	e	4	S	9	7	80	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29

q ò

probable chlorogui probable chlorogui		hypothetical proce hypothetical prote probable membrane	hypothetical prote type I restriction P-type cation tran	hypothetical prote i; hypothetical prote	hypothetical prote hypothetical prote
			:		÷-
T09079 T09080	T18504 T18499	T18491 D71606	T28196 H82934 A44396	D71615 C71609	H64387 A64465
777	9000	4000	000	77	77
2819	522 1619	1119	.439 1045 1984	504 558	377 1005
22.7	, , , , , , , , , , , , , , , , , , ,	7.7.	7.0	7.0	o. o.
142.5	142	139.5	139 138.5 137.5	137 137	136.5 136.5
0 11.0	2 m 4 m	36.0	601	3 5	124

## ALIGNMENTS

 RESULT 1 SAZOK1 major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)	ium falciparum)
C;Species: Plasmodium falciparum C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jun-2000	n-2000
C. M. C. C. S. M. C. S. M.; Bone, N.; Hyde, J.E.; Scaife, J.; Certa, U.; Stunn. F.Mackay, M. 3492-3890, 1985	Stunnenberg, H.
 A; Title: Polymorphism of the precursor for the major surface antigens of Plasmodium f A; Reference number: A91030; MUID:86136024	of Plasmodium f
A;Wolecule type: DNA A;Residues: 1-1631 < MAC>	
C; Comment: The merozoite stages of different strains have strain-specific surface ant C; Comment: P. falciparum has three stages: sporozoite, merozoite, and gametocyte. The	ific surface ant qametocyte. The
C; Superfamily: major merozoite surface antigen C: Rewords: glycoprotein: merozoite: surface antigen: tandem repeat: transm	transmembrane pr
F:1-19/Domain: signal sequence #status predicted <sig> F:20-1631/Product: major merozoite surface antiqen #status predicted <wat></wat></sig>	· · · · · · · · · · · · · · · · · · ·
F;67-84/Region: 3-residue repeats (S-G-T/P) F;1614-1631/Domain: membrane anchor #status predicted <mbn></mbn>	
F; 97, 259, 755, 759, 835, 911, 955, 1049, 1156, 1165, 1436, 1563/Binding site: carbohy	site: carbohydrate (Asn
Query Match 97.3%; Score 1914; DB 1; Length 1631; Best Local Similarity 97.9%; Pred. No. 9.6e-97; Matches 369; Conservative 1; Mismatches 5; Indels 2; Gaps	Gaps 1;
Qy 1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE 60	
DD 1255 AVTPSVIDNILSKIENEYEVLKFLAGVYRSLKKQLENNVMTFNVNVKDILNSRENKRE 1314	
 Qy 61 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120	L 120
 DD 1315 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 1374	L 1374
 QY 121 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180	
 Db 1375 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 1434	, V 1434
 Oy 181 LNYTYEKSNVEIKELIYLKTIODKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 238	
Db 1435 LNYTYEKSNVEVKIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLITKFLSTGM 1494	M 1494
Qy 239 VFENLLKSILSNILD#KLARYVKHFTTP#RKKTMIQQNSGCFRHLDEREECKCLLNYKQE 298	
Db 1495 VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE 1554	E 1554

ä

.; m

84

```
the gp195 (MSA-1) gene from Pl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             parasite (Plasmodium falciparum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     major merozoite surface antigen precursor - malaria parasite (Plasmodium falcipa M; Alternate names: 9p195 surface antigen C; Species: Plasmodium falciparum C; Species: Plasmodium falciparum C; Species: Pasmodium falciparum C; Date: 12-Feb-1993 #text_change 09-Jun-2000 C; Accession: S05603; $04850 R; Myler, P.J.

submitted to the EMBL Data Library, April 1989
A; Reference number: S05603
A; Accession: S05603
A; Accession: S05603
A; Molecule type: mRNA
A; Residues: 1-1639 cMYL->
A; Cross-references: EMBL:X15063; NID:99896; PIDN:CAA33163.1; PID:99897
R; Myler, P.J.
Nucleic Acids Res. 17, 5401, 1989
A; Title: Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene france number: S04850; MUID:89345116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1264 AVTPSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 1323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1324 NFKNVLESDLIPYKDLTSSNYVVKDPYKPLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 1383
                                                                                                                                                                                                                                                             85 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 144
                                                                                                                                                                                                                                                                                                                                     121 GYYKILSEKYKSDLDSIKKYINDKOGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    265 VFBNLAKTVLSNLLDGNL-QGMLNISQHQCVKKQCPQNSGCFRHLDEREECKCLLNYKQE 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     324 GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSNF 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                   61 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANÖVL 120
                                                                                                                                                                                                                                                                                                                                                                  LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                299 GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCSSSNF 358
                                                                                                                              1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE
                                                                                                                                                            25 AVTPSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVATFNVNVKDILNSRFNKRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A)Cross-references: EMBL. 115063
A)Cross-references: EMBL. 115063
C; Superfamily: major merozoite surface antigen
C; Kaywords: alycoprotein; merozoite; surface antigen
E): 1-19/Domain: signal sequence #status predicted <SIG>
F; 20-1639/Product: major merozoite surface antigen #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 1639;
                             Length 400;
                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5;
                             DB 2;
                                                                            22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 89.7%; Score 1763.5; DB Best Local Similarity 91.5%; Pred. No. 1.5e-88; Matches 345; Conservative 7; Mismatches 22
                                                    Pred. No. 3e-89;
                                                                            7; Mismatches
                             Score 1763.5;
                          89.7%;
91.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            384 LGISFLLILMEILYSFI 400
                                                                            Matches 345; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Accession: S04850
A; Molecule type: mRNA
A; Residues: 1504-1639 <MYL2>
                                                    Best Local Similarity
                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        805603
                                                                                                                                 ð
                                                                                                                                                                                 g
                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                         probable major surface antigen (83K, 19K, 42K) precursor - malaria parasite (Plasmodium C.Species: Plasmodium falciparum C.Species: Plasmodium falciparum C.Date: 29-Aug-1987 #sequence_revision 29-Aug-1987 #text_change 09-Jun-2000 C.Accession: A24694 R.Holder, A.A.; Lockyer, M.J.; Odink, K.G.; Sandhu, J.S.; Riveros-Moreno, V.; Nicholls, Asture 317, 270-273, 1985 A.Title: Primary structure of the precursor to the three major surface antigens of Plasm A.Reference number: A24594; MUID:86014355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A45545

Major merozoite surface antigen - malaria parasite (Plasmodium falciparum) (fragments)

major merozoite surface antigen - malaria parasite (Plasmodium falciparum) (fragments)

C;Specias: Plasmodium falciparum

C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jun-2000

C;Accession: A45545

Mol. Blochem. Parasitol. 49, 29-33, 1991

A;Reference number: A45545; MuD:92131048

A;Accession: A45545

A;Accession: A45545

A;Accession: A45545

A;Residues: 1-400 < ELAA

A;Residues: 1-400 < ELAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;NOTE: Sequence extracted from NCBI backbone (NCBIN:77612, NCBIP:77621)
C;Superfamily: major merozoite surface antigen
C;Keywords: glycoprotein; merozoite; surface antigen; tandem repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          239 VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 299 GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCSSSNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 1640;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 1901; DB 2;
Pred. No. 4.9e-96;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Residues: 1-1640 <hOL>
C; Superfamily: major merozoite surface antigen
C; Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              96.68;
97.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1624 LGISFLLILMLILYSFI 1640
                          1615 LGISFLLILMLILYSFI 1631
LGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   359 LGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 97.1
Matches 366; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                    A;Accession: A24594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                               셤
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

;

Gaps

```
A; Reference number: A23386
A; Molecule type: DNA
A; Residues: 1-1104 < WEB1>
A; Cross-references: EMBL:X03831
B; Weber, J.L.; Sim, B.K.L.; Lyon, J.A.; Wolff, R.
Nucleic Acids Res. 16, 1206, 1988
A; Title: Merozoite surface protein sequence from the Camp strain of the human malaria A; Reference number: S06361
A; Accession: S06361
A; Accession: S06361
A; Molecule type: DNA
A; Residues: 1104-1726 <WEB2>
A; Cross-references: EMBL:X03831
C; Comment: The merozoite stages of different strains have strain-specific surface ant C; Comment: P. falciparum has three stages: sporozoite, and gametocyte. The C; Comment: P. falciparum has three stages: sporozoite, and gametocyte. The C; Comment: Signal sequence #status predicted <SIG>C; Superfamily: major merozoite surface antigen
C; Sup
                                                                                                                                                                                                                                       the huma
                                                                                                                                                                                                                                                                                                                                                                                                                              the human malaria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              site: carb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  major merozoite surface antigen precursor – malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
                                                                           falciparum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                .,
,
                                                                                                                                                                                                                                       major merozoite surface antigen of
                                                                                       NiAlternate names: 1956 qlycoprotein
C;Species: Plasmodium falciparum
C;Species: Plasmodium falciparum
C;Accession: A23386; 806361
R;Weber, J.L.; Leininger, W.M.; Lyon, J.A.
Nucleic Acids Res. 14, 3311.332, 1986
A;Fitle: Variation in the gene encoding a major merozoite surface antigen of A;Reference number: A23386; MUID:86205236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            69 DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYP 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9 NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES 68
                                                                         (Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               169 IDLFVIHLEAKVLNYTYEK -- SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              129 KYKSDLDSIKKYI-------NDKQGENEKYLPFLNNIETLYKTVNDK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NNLLTKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             56.4%; Score 1110; DB 1; Length 1726;
56.6%; Pred. No. 6.1e-53;
Live 60; Mismatches 85; Indels 24
                                                                         - malaria parasite
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LFDGIFCSSSNFLGISFLLILMLILYSFI 1726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LSMVIFCSSSNFLGISFLLILMLILYSFI 375
                                                                           precursor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 56.69
Matches 220; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A45948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ŧ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               merc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            parasite (Plasmodium falciparum) (strain RO-71)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4
                             1563 GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSNF 1622
  GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                      239 VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DMLGKLLSTGLV-QNFPNTISKLIEGKF-QDMLNISQHQCVKKQCPQNSGCFRHLDERE 563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               170 DLFVIHLEAKVLNYTYEK--SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHN 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NLLTKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDERE 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPL 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9 NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                              GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCSSSNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     129 KYKSDLDSIKKYI-------NDKQGENEKYLPFLNNIETLYKTVNDKI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 2; Length 651;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    85;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   56.6%; Score 1113.5; DB 57.0%; Pred. No. 1.3e-53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  59; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SMVIFCSSSNFLGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             624 FDGIFCSSSNFLGISFLLILMLILYSFI 651
                                                                                                                                                                                                                                                                                                                                                                                               1623 LGISFLLILMLILYSFI 1639
                                                                                                                                                                                                                                                                                                                                                                           LGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity
les 221; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Loca
Matches
121
                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                           359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           288
ò
                                       g
                                                                                       ά
                                                                                                                            8
                                                                                                                                                                               õ
                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
C. Accession: A26868

R. Tanabe, K.; Mackay, M.; Goman, M.; Scaife, J.G.
J. Moll. Biol. 195, 273-287, 1987

A; Title: Allelic dimorphism in a surface antigen gene of the malaria parasite Plasmod A; Reference number: A26868; MUID:88011243

A; Reference number: A26868; MUID:88011243

A; Recession: A26868

A; Residues: 1-1701 cTAN>
C; Superfamily: major merozoite surface antigen
C; Superfamily: major merozoite surface antigen
F; 1-19/Pomain: signal sequence #status predicted <SIG>
F; 20-1701/Product: major merozoite surface antigen
F; 20-1701/Product: major merozoite surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1611 EREECKCLIAVKQEGDKCEENPAPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPDS 1670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1613 EECKCLLAYKQEGDKCVENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPDSYP 1672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C;Species: Plasmodium falciparum
C;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 09-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNLNDILNSRLKKRKYFLDVLES 1374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYP 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    169 IDLFVIHLEAKVLNYTYEK -- SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNH 226
                                                                                                         DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----NDKQGENEKYLPFLNNIETLYKTVNDK
                                                                                                                                                                                                         KYKSDLDSIKKYI---------NDKQGENEKYLPFLNNIETLYKTVNDK
                                                                                                                                                                                                                                                                                                                                           IDLFVIHLEAKVLNYTYEK -- SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NNLLTKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9 NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NNLLTKFLSTGMV - - FENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EREECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 1701;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match 54.8%; Score 1077; DB 2; Best Local Similarity 55.5%; Pred. No. 3.8e-51; Matches 217; Conservative 60; Mismatches 86;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         347 LSMVIFCSSSNFLGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KYKSDLDSIKKYI------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       σ
          1315
                                                                                                                                                                                                         129
                                                                                                                                                                                                                                                                       1435
                                                                                                                                                                                                                                                                                                                                           169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        285
                                                                           69
                                                                                                                                     1375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           69
                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                    δ
                                                                                                                                     g
                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A54498
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (is C;Species: Plasmodium falciparum
C;Species: Plasmodium falciparum
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C;Accession: A54498
R;Peterson, M.G.; Coppel, R.L.; McIntyre, P.; Langford, C.J.; Woodrow, G.; Brown, G.V.; Mol. Blochem. Parasitol. 27, 291-302, 1988
A;Title: Variation in the precursor to the major merozoite surface antigens of Plasmodiu A;Reference number: A54498; MuID:88142999
                                                                                                                                               meroz
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jun-2000 C;Accession: A4548 B;Chang, S.P.; Kramer, K.J.; Yamaga, K.M.; Kato, A.; Case, S.E.; Siddiqui, W.A. Exp. Parasitol. 67, 1-11, 1988 A;Title: Plasmodium falciparum: gene structure and hydropathy profile of the major A;Accession: A45948; MUID:89005525 A;Accession: A45948 A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 227 NNLLTKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDER 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYP 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9 NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NID:g160412; PIDN:AAA29653.1; PID:g160413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               169 IDLFVIHLEAKVLNYTYEK--SNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    129 KYKSDLDSIKKYI--------NDKQGENEKYLPFLNNIETLYKTVNDK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 1726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 1701;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      87; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        56.1%; Score 1103; DB 2; 56.3%; Pred. No. 1.5e-52;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Score 1091; DB 2;
; Pred. No. 6.5e-52;
61; Mismatches 87;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         60; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1701 <PET>
A;Cross-references: GB:M19143; NID:g160412; PIIC
C;Superfamily: major merozoite surface antigen
C;Reywords: surface antigen
                                                                                                                                                                                                                                                                                                                                    A:Cross-references: GB:M37213
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LSMVIFCSSSNFLGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        55.5%;
55.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 56.3
Matches 219; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 55.55
Best Local Similarity 55.87
Matches 217; Conservative
                                                                                                                                                                                                                                                                                                  A; Residues: 1-1726 <CHA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
```

g

ð

287

ò g õ

g

à

셤

69

ò

a

ò g οy

g

ð

S

.;

ð

```
C; Species: Plasmodium yoelii
C; Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 26-Aug-1999
C; Accession: A28121
R; Burns Jr., J.M.; Dally, T.M.; Vaidya, A.B.; Long, C.A.
R; Burns Jr., J.M.; Dally, T.M.; Vaidya, A.B.; Long, C.A.
Proc. Natl. Acad. Sci. U.S.A. 85, 602-606, 1988
A; Title: The 3' portion of the gene for a Plasmodium yoelii merozoite surface antigen
A; Reference number: A28121; MUID:88124889
                                                                                                                                                                                                                                                                         1604 VKSSGLLEKLMKSKLIKENESKEILSQLLNVQTQLLTMSSEHTCI----DTNVPDNAACY 1659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1491 GVKTEIKKVEDDIKKQDEELKKLGNVNSQDSKKNEFIAKKAELEKYLPFLNSLQKEYESL 1550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     76 LTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSEKYKSDLD 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -- KQGENEKYLPFLNNIETLYKTV 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      166 NDKIDLFVIHLEAKVLNYTYEKSNVEI--KELIYLKTIQDKLADFKKNNNFVGIADLSTD 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     :|||: || :| || || ||-||||| :
DILSEFTNESLYYYTKRLGSTYKSLKKHMLREFSTIKEDMTNGLNNKSOKRNDFLEVLSH 368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KYKSDLDSIKKYI------ND-KQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                             16 NEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLESDLIPYKD 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9 NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES 68
  sequence extracted from NCBI backbone (NCBIN:83591, NCBIP:83592)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RHLDEREECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             224 YNHNNLLTKFLSTGMVFENLLKSILSNLLDWK---LARYVKHFTTPMRKKTMIQQNSGCF
                                                                                                                                                                                           50;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A.Accession: A28121
A.Molecule type: DNA
A.Residues: 1-680 cSUR>
A.Cross-references: GB:303612; NID:g160678; PID:g160679
A.Cross-references: GB:403612; NID:g160678; PID:g160679
A.Experimental Source: strain 17XL
A.Note: the authors translated the codon GTA for residue 429 as Leu
C.Superfamily: major merczotte surface antigen
C.Superfamily: major merczotte
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 680;
                                                                                                                                      Length 1751;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          major merozoite surface antigen - Plasmodium yoelii (fragment)
                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                   Query Match 36.2%; Score 713; DB 2; L
Best Local Similarity 39.1%; Pred. No. 2.6e-31;
Matches 153; Conservative 71; Mismatches 117;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28.7%; Score 565.5; DB 2; 34.9%; Pred. No. 9.4e-24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    73; Mismatches 144;
A;Note: sequence extracted from NCBI backbone C;Superfamily: major merozoite surface antigen C;Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1716 KEGSEPLFEGVFCSSSSFLSLSFLLLMLLFL 1746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 341 KPDSYPLSMVIFCSSSNFLGISFLLILMLIL 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 34.9
Matches 134; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIK --- KYIND---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        129
                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       С
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                         merozoite surface antigen 1 precursor - Plasmodium vivax
C;Species: Plasmodium vivax
C;Species: Plasmodium vivax
C;Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 24-Nov-1999
C;Accession: A39401
R;del Portillo, H.A.; Longacre, S.; Khouri, E.; David, P.H.
Proc. Natl. Acad. Sci. U.S.A. 88, 4030-4034, 1991
A;Title: Primary structure of the merozoite surface antigen 1 of Plasmodium vivax reveal
A;Reference number: A39401; MUID:91219506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A45604

major blood-stage surface antigen Pv200 - Plasmodium vivax

c;Species: Plasmodium vivax

C;Species: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999

C;Accession: A45604

K;Glbson, H.L.; Tucker, J.E.; Kaslow, D.C.; Krettli, A.U.; Collins, W.E.; Kiefer, M.C.;

Mol. Biochem. Parasitol. 50, 325-333, 1992

A;Title: Structure and expression of the gene for Pv200, a major blood-stage surface ant A;Reference number: A45604; MUID:92158013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1466 AVNEEVKKVEADIKAEDDKIKKIGSDSTKTTEKTQSMAKKAELEKYLPFLNSLQKEYESL 1525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----DSIKKYIND------KQGENEKYLPFLNNIETLYKTV 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            76 LTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSEKYKSDL- 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           166 NDKIDLFVIHLEAKVLNYTYEKSNVEI -- KELIYLKTIQDKLADFKKNNNFVGIADLSTD 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16 NEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLESDLIPYKD 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     224 YNHNNLLTKFLSTGMVFENLLKSILSNLLDWK---LARYVKHFTTPMRKKTMIQQNSGCF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               281 RHLDEREECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  50;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 1726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          36.3%; Score 714; DB 2; Length 17
39.6%; Pred. No. 2.2e-31;
tive 67; Mismatches 119; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: A45604
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-1751 < CLB>
A; Cross-references: GB:M75674; NID:g160608; PID:g457336
                                                  1671 YPLFDGIFCSSSNFLGISFLLILMLILYSFI 1701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Superfamily: major merozoite surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     341 KPDSYPLSMVIFCSSSNFLGISFLLILMLIL 371
                         345 YPLSMVIFCSSSNFLGISFLLILMLILYSFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 39.69
Matches 155; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Residues: 1-1726 <DEL>
A; Cross-references: GB:M60807
C; Superfamily: major merozolte
C; Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                     g
```

g

ò

g

ò

qq

Q ò

ò

οž

g

ò

οy

œ

ဖ

```
C; Species: Plasmodium chabaudi chabaudi
C; Species: 91-891-1993 #sequence_revision 30-Sep-1993 #text_change 26-Aug-1999
C; Datcession: A45546
R; Delecrahijder, W.; Hendrix, D.; Bendahman, N.; Hanegreefs, J.; Brijs, L.; Hamers-Ca Mol. Blochem: Parasitol. 43, 231-244, 1990
A; Title: Molecular cloning and sequence analysis of the gene encoding the major meroz A; Reference number: A45546; MUID:91218805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORF MSV152 probable core protein P4a homolog (vaccinia A10L) . Melanoplus sanguinipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Cross-references: EMBL:AF063866; NID:94049647; PIDN:AAC97675.1; PID:94049715
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C;Species: Melanoplus sanguinipes entomopoxvirus
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1637 NNLNKSGLVTEGESKKILAKMINMDAMDLLGIGSNHVCISTSTPDNAGCFRYDDGTEEWR 1696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1529 FNKMIELYKPQLNAVNEQIAAIGTEPTDAEKKKYAPIFEDLKGLYETILNGAEEFSELLQ 1588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   291 CLLNYKQ--EGSKCVENPNPTCNENNGGCDADAKCTEEDSGSN--GKKITCQCTKPDSYP 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63 KNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGY 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 123 YKILSEKYKSDLDSIKKYI-----NDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLE 177
                                                                                                                                                               major merozoite surface antigen precursor - Plasmodium chabaudi chabaudi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 TPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      231 TKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   D.L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             178 AKVLNYTYEKSNVEIKELIYLKT --- IQDKLADF ----KKNNNFVGIADLSTDYNHNNLL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C,Accession: T28313
R;Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, J. Virol. 73, 533-552, 1999
A;Title: The genome of Melanoplus sanguinipes entomopoxvirus.
A;Reference number: Z20484; MUID:99102612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 1785;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Cross-references: GB:M34947; NID:g160597; PID:g160598 C;Superfamily: major merozoite surface antigen C;Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26.9%; Score 529; DB 2; I
31.5%; Pred. No. 2.8e-21;
iive 79; Mismatches 158;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Status: preliminary; translated from GB/EMBL/DDBJ A;Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1757 YYAGVECSSSGFMGLSILLIITLIVFN 1783
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  347 LSMVIFCSSSNFLGISFLLILMLILYS 373
                        1747 GVFCSSSFMGLSILLIITLIVFN 1770
350 VIFCSSSNFLGISFLLILMLILYS 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 31.59
Matches 122; Conservative
                                                                                                                                                                                                                                                                                                                                                                              A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1785 <DEL>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Residues: 1-1306 <AFO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Accession: T28313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                     14
                                                                                                                     RESULT
                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                           A45532
major merozoite surface antigen precursor - Plasmodium yoelii
C;Species: Plasmodium yoelii
C;Species: Plasmodium yoelii
C;Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 26-Aug-1999
C;Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 26-Aug-1999
C;Accession: A45531
R;Lewis, A.P.
Mol. Blochem. Parasitol. 36, 271-282, 1989
A;Title: Cloning and analysis of the gene encoding the 230-kilodalton merozoite surface
A;Reference number: A45532; MUID:90014981
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Status: preliminary
A; Molecule type: DNA
A; Molecule type: J-772 <LEW>
A; Cross-references: GB.J04668; NID:g160492; PID:g160493
B; Daly, T.M.; Burns Jr., J.M.; Long, C.A.
R; Daly, T.M.; Barns Jr., J.M.; Long, C.A.
A; Title: Percursor to the major merozoite surface antigen of Plasmodium yoelii: cloning
A; Reference number: A45531; MUID:90014982
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DLIPYKDLISSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKILSE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KYKSDLDSIKKYI-----ND-KQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                           CLLNYKQ-EGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSM 349
181 LNYTYEKSNVEI--KELIYLKTIQDKLADF----KKNNNFVGIADLSTDYNHNNLLTKFL 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                       NILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LNYTYEKSNVEI -- KELIYLKTIQDKLADF -- -- KKNNNFVGIADLSTDYNHNNLLTKFL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              291 CLLNYKQ-EGSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSM
                                                                                            235 STGMVFENLLKSILSNLLDWK----LARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STGMVFENLLKSILSNLLDWK----LARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 1772;
                                          489 DNYKNEKAEFEILTKNLEKYIQIDEKLDEFVEHAENNKHIASIA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Status: preliminary
A;Molecule type: DNA
A;Residues: 454-1094 < CDAL>
A;Cross-references: GB:J03975; NID:g160081; PID:g160082
C;Superfamily: major merozoite surface antigen
C;Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 28.7%; Score 565.5; DB 2; Best Local Similarity 34.9%; Pred. No. 2.8e-23; Matches 134; Conservative 73; Mismatches 144;
                                                                                                                                                                                                                                                                                  350 VIFCSSSNFLGISFLLILMLILYS 373
                                                                                                                                                                                                                                                                                                             655 GVFCSSSSFMGLSILLITTLIVFN 678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Accession:
                                                                                                                                                                                                                                                                                                                                                                                                          13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  235
                                                                                                                                                                                        291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                     a
                                                                                                                                 셤
                                                                                                                                                                                      ò
                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
ð
                                                                                            à
```

<u>بر</u>

C;Genetics: A;Note: MSV152

13; 150 -----KYLPFLNNIETLYKTVNDKIDLEVIHLEAK-VLNYTYEKSNVEIKELIYLKTIQ 202 :|:| :: | :: | :: | :: | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 -----LESDLIPYKDLTSSNYVVKDPYKF---LNKEKRDKFLSSYNYIKDSIDTDINFA 116 -----FVGIADLSTDYNHNNLLTKFLS--TGMVF 240 :| : || || || :: 303 LRLYENKKYYNNNNALSKFVSNKLLDIQRRIINIMEINSDNNPTEILITFNSKYQSLII 362 Gaps 8 HNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNV-- 65 Query Match 8.4%; Score 165.5; DB 2; Length 1306; Best Local Similarity 21.1%; Pred. No. 0.12; Matches 72; Conservative 63; Mismatches 111; Indels 95; 241 ENLLKSIL--SNLLDWKLARYVKHFTTPMRKKTMIQQNSGC 279 117 NDVLGYYKILSEKYKSDL----DSIKKYINDKQGENE-----203 DKLADFKK----NNN-----99 g g ò à ò g ò g δ g ò

Search completed: August 8, 2001, 12:30:58 Job time: 106 sec

Sequence 15, Appl Sequence 6, Appli Sequence 184, Appli Sequence 2, Appli

Sequence 1, Sequence 4,

ALIGNMENTS

Sequence 1, A Sequence 20, Sequence 20, Sequence 15,

Run on:

```
GENERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.
APPLICANT: ELACKEN, MICHAEL J.
APPLICANT: ELACKEN, MICHAEL J.
APPLICANT: ELACKEN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPORTION: NACCINE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINTON
CONTRY: O.S.A.
ZIP: APPLICATION NOW PC COMPATIBLE
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER READABLE FORM:
MEDIUM TYPE: Ploppy disk
COMPUTER READABLE FORM:
MEDIUM TYPE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919
FILING DATE: 22-FEB-1992
PRICE APPLICATION DATE: APPLICATION NUMBER: 16,773
REPERENCE/CATION NUMBER: 16,773
RECISTRATION NUMBER: 16,773
RECISTRATION NUMBER: 16,773
RECISTRATION NUMBER: 16,773
RECIEDENOR: (202) #61-3000
TELEPHONE: (202) #61-300
TELEPHONE: CHARACTERICS:
LINGTHH: 53 amino acids
                                   US-08-466-961A-20
US-08-645-193B-15
US-07-798-915A-6
US-08-005-002C-6
US-08-481-105-6
US-08-484-105-6
US-08-929-329-6
US-08-929-329-6
US-08-928-865A-2
US-09-98-865A-2
US-09-9323-735-2
                                                                                                                                                                                               US-09-080-897-4
US-08-899-595-1
                                                                                                                                                                                                                                                                                                                              Sequence 3, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME/KEY: Modifled-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS: single
720
720
990
990
990
1264
1264
615
615
615
730
1248
1258
1258
1255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY: linear MOLECULE TYPE: protein FEATURE:
 US-08-290-919-3
108.5
108.5
108.5
107.5
107.5
107.5
105.5
105.5
105.5
103.5
103.5
103.5
103.5
103.5
103.5
103.5
103.5
103.5
103.5
103.5
103.5
(without alignments)
224.393 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 3, Appli
Sequence 11, Appl
Sequence 11, Appl
Sequence 12, Appl
Sequence 2, Appli
Sequence 2, Appli
Sequence 3, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Appli
Appli
Appli
Appli
Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sequence 2, Ap
Sequence 120,
Sequence 120,
Sequence 2, Ap
Sequence 2, Ap
Sequence 21, A
Sequence 21, A
Sequence 21, A
Sequence 21, A
Sequence 5, Ap
Sequence 5, Ap
Sequence 6, Ap
Sequence 2, Ap
                                                                                                                                                                  1 AVTPSVIHNILSKIENEYEV.....SNFLGISFLLILMLILYSFI 37
                                                                                       ; Search time 34.41 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                                                                                                                                                                                             /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
            GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-290-919-3

US-08-290-919-4

US-08-290-919-11

US-08-290-919-12

US-08-290-919-12

US-08-290-919-12

US-08-290-919-12

US-08-290-919-12

US-08-257-07-3

US-08-257-07-3

US-08-46-855A-2

US-08-46-855A-2

US-08-46-855A-2

US-08-46-855A-2

US-08-46-85A-2

US-08-418-2

US-08-418-2

US-08-411-5

US-08-411-5

US-08-411-5

US-08-411-5

US-08-411-5

US-08-411-5
                                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08-480-604A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10-551B-2
                                                                                                                                                                                                                                    197339 seqs, 20590346 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUMMARIES
                                                                                                                                                                                                                                                                                                                                                        Listing first 45 summaries
                                                              OM protein - protein search, using sw model
                                                                                      August 8, 2001, 12:29:58
                                                                                                                                                                                              BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0% Maximum Match 100%
                                                                                                                                                                                                                                                                                                                                                                                   Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                     Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                          US-09-500-376-4
1967
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           48
984
984
23391
23391
3135
1786
455
455
455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11147
11147
22710
3289
912
599
1435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   112.5
112.5
112.5
112.5
112.5
112.5
112.5
112.5
110.5
                                                                                                                                                       Perfect score:
                                                                                                                                                                                              Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score
                                                                                                                                                                       Sequence:
                                                                                                                                                                                                                                     Searched:
                                                                                                                                                                                                                                                                                                                                                                                   Database
```

Result No.

ô

Gaps

ö

Indels

```
306 PNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCSSSN 357
                                                                 2 PNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPDSYPLFDGIFCSSSN 53
    Mismatches
                                                                                                                                                                                  Sequence 11, Application US/08290919
Patent No. 5720959
    ;;
;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       106 amino acids
    45; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
Matches 47; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: amino acid
STRANDEDNESS: si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                               US-08-290-919-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-290-919-11
    Matches
                                         ò
                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                 ö
                                                                                                                                                                   Gaps
                                                                                                                                                                 ö
                                                                                                                                                                                                       306 PNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCSSSN 357
                                                                                                                                                                                                                          2 PNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSN 53
                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
TITLE OF INVENTION: YACCINE
NUMBER OF SEQUENCES: 19
                                                                                                                         Length 53;
                                                                                                                                                               3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CLASSIFICATION: 435
PRIOR APPLICATION 1435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CB93/00367
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/6BC8/
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
ELLOALIOATION NUMBER: US/08/290,919
FILING DATE: 04-OCT-1994
                    /label= X = M and N, or N"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 - Mand N, or N"
                                                                                                                     Score 272; DB 1;
Pred. No. 2.9e-15;
1; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P. STREET: 1100 NEW YORK AVENUE, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                             Sequence 4, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /label= X
/note= "X
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
                                                                                                                       13.8%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME/KEY: Modified-site
CCCATION: 1
COTHER INFORMATION: /labe
CTHER INFORMATION: /note
US-08-290-919-4
                                                                                                             Query Match
Best Local Similarity 92.38
Matches 48; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN D
            CTHER INFORMATION:
CTHER INFORMATION:
US-08-290-919-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY: WASHINGTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: U.S.A. ZIP: 20005-3918
LOCATION: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
                                                                                                                                                                                                                                                                                                                          US-08-290-919-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STATE:
                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                          g
```

```
5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       261 KHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQ-EGSKCVENPNPTCNENNGGCDA 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               12 KHVCVDTRD---IPKNAGCFRDDNGTEEWRCLLGYKKGEGNTCVENNNPTCDINNGGCDP 68
Jequeince
Patent No. 572095,
GENERAL INFORMATION:
APPLICANT: HOLDER; ANTHONY A.
APPLICANT: BLACKMAN, MICHARL J.
APPLICANT: CHAPPEL, JONATHAN L.
APPLICANT: CHAPPEL, AND L.
APPLICA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 106;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CURRENT APPLICATION DATA:
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919
FILING DATE: U-CCT-1994
CLASSIFICATION 35
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REGISTRATION NUMBER: 212242/HCM/MJL/6BC8/
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEPHONE: (202) 822-0944
TELEX: 6714627 CUSH
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   69 TASCONAESTENSKKIICTCKEPTPNAYYEGVFCSSSS 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11.8%; Score 232; DB 1; ilarity 48.0%; Pred. No. 1.1e-11; Conservative 11; Mismatches 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        320 DAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCSSSN 357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY: WASHINGTON STREET: 1100 NEW YORK AVENUE, N.W. CITY: WASHINGTON STRATE: D.C. STRATE: D.C. ZIE: 20005-3918 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
```

US-08-290-919-1

Score 253; DB 1; Length 53; Pred. No. 9.3e-14;

12.9%; 86.5%;

Query Match Best Local Similarity

```
Patentin Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 200; DB 1;
Pred. No. 1.4e-09;
0; Mismatches 5
                                          E: CUSHMAN DARBY & CUSHMAN, L.L.P. 1100 NEW YORK AVENUE, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STREET: 1100 NEW YORK AVENUE, N.W. CITY: WASHINGTON STATE: D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPALIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 87.5%;
Matches 35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 1
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 48 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             , MOLECULE TYPE: peptide US-08-290-919-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
                         CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        amino acid
                                                                                   CITY: WASHINGTON
                                                                                                      STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RY: U.S.A.
20005-3918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRANDEDNESS:
                                            ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JS-08-290-919-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 12, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA TITLE OF INVENTION: VACCINE
                                                                            APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 200; DB 1; Length 48;
Pred. No. 1.4e-09;
0; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                            COMPUTER: IEM PC compatible
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919
FILING DATE: 04-0CT-1994
CLASSITCATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 212242/HCM/MJL/6BC8/
TELECOMMUNICATION INFORMATION:
TELEPRAX: (202) 861-3000
TELEFRAX: (202) 861-3000
TELEFRAX: (202) 822-0944
SIGNICATION FOR SEQ ID NO: 1:
SEQUINCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                270 KTMIQQNSGCFRHLDEREECKCLLNYKQEGSKCVENPNPT 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ) OTHER INFORMATION: /label= X OTHER INFORMATION: /note= "X = M and N, or N" US-08-290-919-1
                                                                                                                                                                                                    CUSHMAN DARBY & CUSHMAN, L.L.P.
                                                                                                                                                                                                                     STREET: 1100 NEW YORK AVENUE, N.W.
Sequence 1, Application US/08290919
Patent No. 5720959
                                                           HOLDER, ANTHONY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 10.2%;
Best Local Similarity 87.5%;
Matches 35; Conservative
                                                                                                                                                                                                                                                                                                                                                 Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME/KEY: Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 48 amino acids
                                                                                                                                                                                                                                                                                   COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: protein
                                                                                                                                                                                                                                             CITY: WASHINGTON STATE: D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linear
                 GENERAL INFORMATION:
APPLICANT: HOLDER,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRANDEDNESS:
                                                                                                                                                                                                      ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 5
US-08-290-919-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LOCATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                          ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 2, Application US/08290919
Sequence 2, Application US/08290919
GENERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKWAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
                                                                                                                                                                                                                                                                                                                                                                    Length 48;
                                                                                                                                                                                                                                                                                                                                                                                          Indels
270 KTMIQQNSGCFRHLDEREECKCLLNYKQEGSKCVENPNPT 309
```

ö

```
155 LINIET----LYKT--VNDKIDLFVIH-----LEAKVLNYTYEKSNVEIKELIYLKTIQD 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             204 KLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGMVFENL-----LKSILSNLLDWKLA 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         509 YFNEHLYNDKNSPEDKDNKGKGVVHVDTTLEKEDTLSYDNSDNMFCNKEYCNRLKDENNC 568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 236 PRNIQNICETGKNFKLVVYIKENTLIIKWKVYGETKDTTENNKVDVRKYLINEKE---TP 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  293 FTSILIHAYKEHNGTNLIESKNYALGSDIPEKCDTLASNCFLSGNFNIEKCFQCALLVEK 352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     95 RDKFLSSYNYIKDSIDTDINFANDVLGYYKILSEKYKSDLDSIKKYINDKQGENEKYLPF 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----NDVC--YKYLSEDIVSKFKEIKAETEDDDEDDYTEYKL 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               280 FRHLDEREECKC-----LLNYKQEGSKCVENPNPT-------C--NENNGGCDADAK 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 PSVIHNILSKIENEYEVLYLKPLA----GVYRSLKKQLENN---VMTFNVNVKDILNSP 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    56 FN------KRENFKNVLES------DLIPYK-DLTSSNYVVKDPYKF-----LNKEK 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           258 RYVKHF------GC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 6.5%; Score 127; DB 1; Length 984; Best Local Similarity 20.4%; Pred. No. 0.051; Matches 94; Conservative 64; Mismatches 161; Indels 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RECOMBINANT VIRUS IMMUNOTHERAPY 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          323 CTE----EDSG-----SNGKKITCQCTKPDSYPLSM 349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      629 PMEFLQIIEDYGFLPAESNYPYNYVKVGEQCPKVEDHWMNL 669
                                                                                                                  NAME: Fronmer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2570
TELECOMMUNICATION INFORMATION:
TELEPAX: (212) 840-333
TELERX: (212) 840-0712
TELEX: 42506 CURTMS.
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Pacletti, Enzo
APPLICANT: Tartaglia, James
APPLICANT: Cox, William I.
TITLE OF INVENTON: RECOMBINANT VIRUS IM
NUMBER OF SEQUENCES: 217
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
                                             US 07/672,183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 120, Application US/08184009
Patent No. 5833975
GENERAL INFORMATION:
18-MAR-1992
                                                                                         ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S
                                                                  20-MAR-1991
                                                                                                                                                                                                                                                                                                                                              LENGTH: 984 amino acids
TYPE: amino acid
STRANDEDNESS: single
                      PRIOR APPLICATION DATA:
APPLICATION NUMBER: U
FILING DATE: 20-MAR-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   :|
353 ENK-----
                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: lin
MOLECULE TYPE:
FRAGMENT TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-184-009-120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-257-073-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 197; DB 1; Length 48;
Pred. No. 2.4e-09;
1; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
                                                                                    PRIOR PAPLICATION DATA:
APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/6BC8/
TELECOMMUNICATION INFORMATION:
TELECAMMUNICATION INFORMATION:
TELECAMMICATION INFORMATION:
TELECAMMICATION INFORMATION:
TELECAMMICATION INFORMATION:
TELECAMMICATION INFORMATION:
TELECAMMICATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   270 KTMIQQNSGCFRHLDEREECKCLLNYKQEGSKCVENPNPT 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /label= X
/note= "X = M and N, or N"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE: Curtis, Morris & Safford, P.C. STREET: 530 Fifth Avenue, 25th Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: US/08/257,073
FILING DATE: 09-JUN-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/075,783
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
                    APPLICATION NUMBER: US/08/290,919
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             UNITED STATES OF AMERICA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: US 07/852,305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 3, Application US/08257073
Patent No. 5766597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Paoletti, Enzo
APPLICANT: de Taisne, Charles
APPLICANT: Tine, John A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10.0%;
85.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME/KEY: Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGIH: 48 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 10.0 Best Local Similarity 85.0 Matches 34; Conservative
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; LOCATION: 1
; OTHER INFORMATION:
; OTHER INFORMATION:
US-08-290-919-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STATE: New York COUNTRY: UNITED (21P: 10036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
                                           FILING DATE: 04
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-257-073-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

```
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       293 FTSILIHAYKEHNGTNLIESKNYALGSDIPEKCDTLASNCFLSGNFNIEKCFQCALLVEK 352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       95 RDKFLSSYNYIKDSIDTDINFANDVLGYYKILSEKYKSDLDSIKKYINDKQGENEKYLPF 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          155 LNNIET----LYKT--VNDKIDLFVIH-----LEAKVLNYTYEKSNVEIKELIYLKTIQD 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  392 TESIDNILVKMFKTNENNDKSELIKLEEVDDSLKLELMNYCSLLKDVDTTGTLDNYGMGN 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             204 KLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGMVFENL-----LKSILSNLLDWKLA 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               509 YFNEHLYNDKNSPEDKDNKGKGVVHVDTTLEKEDTLSYDNSDNMFCNKEYCNRLKDENNC 568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              280 FRHLDEREECKC-----LLNYKQEGSKCVENPNPT-------C--NENNGGCDADAK 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :| :: | | : | | : | | : | | : | | : | | : | | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            353 ENK---------NDVC--YKYLSEDIVSKFKEIKAETEDDDEDDYTEYKL 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FN-----KRENFKNVLES-----DLIPYK-DLTSSNYVVKDPYKF-----LNKEK 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4 PSVIHNILSKIENEYEVLYLKPLA----GVYRSLKKQLENN---VMTFNVNVKDILNSP 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           258 RYVKHF-----GC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 6.5%; Score 127; DB 2; Length 984; Best Local Similarity 20.4%; Pred. No. 0.051; Matches 94; Conservative 64; Mismatches 161; Indels 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             323 CTE----EDSG----SNGKKITCQCTKPDSYPLSM 349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               629 PMEFLQIIEDYGFLPAESNYPYNYVKVGEQCPKVEDHWMNL 669
                                                                                                                                                                                                                                                                                                                                                   SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/184,009
                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: US/US/184, UU3
FILING DATE: 19-JAN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
TELEROMMUNICATION INFORMATION:
TELEROME: (212) 840-333
TELERAX: 42506GCUTMS
TELERX: 42506GCUTMS
INFORMATION FOR SEQ ID NO: 120:
SEQUENCE CHARACTERISTICS:
LENGTH: 984 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                    MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
                                                                                                                                                                              COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
New York
                                                                                                                               10036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-184-009-120
                                                                                         COUNTRY:
CITY:
STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ζ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

Sequence 120, Application US/08458356 Patent No. 5942235

US-08-458-356-120

```
293 FTSILIHAYKEHNGTNLIESKNYALGSDIPEKCDTLASNCFLSGNFNIEKCFQCALLVEK 352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         95 RDKFLSSYNYIKDSIDTDINFANDVLGYYKILSEKYKSDLDSIKKYINDKQGENEKYLPF 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               : :| | | SOG YFNEHLYNDKNSPEDKDNKGKGVVHVDTTLEKEDTLSYDNSDNMFCNKEYCNRLKDENNC 568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         155 LNNIET----LYKT--VNDKIDLFVIH-----LEAKVLNYTYEKSNVEIKELIYLKTIQD 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :|: ::|| ||| :| :::|| || :| :::|| || :|::|| :|:|| :|::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|:::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|::|| :|:|| :|::|| :|::|| :|:|| :|::|| :|:|| :|::|| :|::|| :|:|| :|:|| :|:|| :|::|| :|
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4 PSVIHNILSKIENEYEVLYLKPLA----GVYRSLKKQLENN---VMTFNVNVKDILNSP 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 56 FN-----KRENFKNVLES-----DLIPYK-DLTSSNYVVKDPYKF-----LNKEK 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               204 KLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGMVFENL-----LKSILSNLLDWKLA 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           280 FRHLDEREECKC-----LLNYKQEGSKCVENPNPT------C--NENNGGCDADAK 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 6.5%; Score 127; DB 2; Length 984; Best Local Similarity 20.4%; Pred. No. 0.051; Matches 94; Conservative 64; Mismatches 161; Indels 14
APPLICANT: Paoletti, Enzo
APPLICANT: Tartaqlia, James
APPLICANT: Cox, William I.
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
NUMBER OF SEQUENCES: 217
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---TTPMRKKTMIQQNS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,356
FILING DATE: 02-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/184,009
FILING DATE: 19-JAN-1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME: Frommer, William S.
REGISTATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2530
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                           ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 42506GURTMS
INFORMATION FOR SEQ ID NO: 120:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 984 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRANDEDNESS: SINGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE:
FRAGMENT TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             258 RYVKHF----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-458-356-120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
```

ဖ

셤 ò 셤

```
| : : | | : : | : : | : : | : : | | : : | : : | : : | : : | | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ||:| || || : ::| || || 143 NSVINNKNKENSSLLK------GDEEDIVMVNLK-KENNYNSVINNVDCRKKDM 1489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1392 DINILSAQNSNNNFSCNNENMNKANVDVNVLENDT------KKREDINTTTVFMEGQ 1442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----YLDNFNTSDEEIGNNKNMDMYLSKEKSISNKNPGNSYYVVDSVXNNEYKI 1581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     86 PYKFLNKEKRDKFLSSYNYIKDSIDTDIN-FANDVLGYYKILSEKYKSDLDSIKKY---- 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      141 ---INDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKVLNYTYEKSNVEI-KELI 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               197 YLKTIQDKLADFKK-----NNNFVGIADLSTDYNHNNLLTKFLSTGMVFENLLKSIL 248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --MRKKTMIQQNSG-CFRHLDEREECKCLL 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            37 LENNV----MTFNVNVKDILNSPFNKRENFKNVLESDLIPYKDLTSSNYVV-----KD 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate Patent No. 6183996
TITLE OF INVENTION: Synthetase II
FILE REFERENCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 2391;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; OTHER INFORMATION: Description of Artificial Sequence: protein US-09-150-741-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                            1582 NKMKE---LIDNENLNDEYNN---NVNMNCSNYNNASAFVNGK 1618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         294 NYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGS---NGK 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             294 NYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSGS---NGK 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match 6.0%; Score 117.5; DB 4; Best Local Similarity 20.1%; Pred. No. 0.95; Matches 69; Conservative 60; Mismatches 111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CURRENT FILING DATE: 1998-09-10
EARLIER APPLICATION NUMBER: PL6380
EARLIER FILING DATE: 1992-12-16
EARLIER APPLICATION NUMBER: AU93/00617
EARLIER FILING DATE: 1993-12-02
EARLIER FILING DATE: 1993-12-05
EARLIER FILING DATE: 1995-07-06
NUMBER OF SEQ ID NOS: 15
SEQ ID NO 2
LENGTH: 2391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CURRENT APPLICATION NUMBER: US/09/150,741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 2, Application US/08323170B Patent No. 573372 GENERAL INFORMATION: APPLICANT: Williamson, Kim C. APPLICANT: Kaslow, David C.
                                                                                                                                                                                                                                                                                                                                                                                      Sequence 2, Application US/09150741 Patent No. 6183996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GENERAL INFORMATION:
APPLICANT: Stewart et al.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SNLLDWKLARYVKHFTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-323-170B-2
                                                                                                                                                                                                                                                                                                                                                         US-09-150-741-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1532
                                            1532
                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     141 ---INDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKVLNYTYEKSNVEI-KELI 196
: | : | | : | | 569 ISNLQVEDQGNCDTSWIFASKYHLETIRCMKGYEPTKISALYVANCYKGEHKDRCDEGSS 628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    86 PYKFLNKEKRDKFLSSYNYIKDSIDTDIN-FANDVLGYYKILSEKYKSDLDSIKKY---- 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              197 YLKTIQDKLADFKK-----NNNFVGIADLSTDYNHNNLLTKFLSTGMVFENLLKSIL 248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SNLLDWKLARYVKHFTTP-------MRKKTMIQQNSG-CFRHLDEREECKCLL 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            37 LENNV---MTFNVNVKDILNSPFNKRENFKNVLESDLIPYKDLTSSNYVV-----KD 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 2391;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Stewart, Thomas S
APPLICANT: Stewart, Thomas S
APPLICANT: Flores, Maria V
APPLICANT: Flores, Maria V
APPLICANT: O'Sullivan, William J
TITLE OF INVENTION: Uncleotide sequence encoding carbamoyl
TITLE OF INVENTION: phosphate synthetase II
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEE: Nixon & Vanderhye PC: 1100 No. 5849573th Glebe Road, 8th Floor Arlington
                                                                                                                        ---SNGKKITCQCTKPDSYPLSM 349
                                                                                                                                                                                             629 PMEFLQIIEDYGFLPAESNYPYNYVKVGEQCPKVEDHWMNL 669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; Score 117.5; DB 2;
; Pred. No. 0.95;
60; Mismatches 111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1490 DGKNINDECKTYKKNKYKDMGLNNNIVDELSNGTSHSTNDHL--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DARA:
APPLICATION NUMBER: US/08/446,855A
FILING DATE: 06-Jul-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                      Sequence 2, Application US/08446855A Patent No. 5849573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29.009
REFERENCE/DOCKET NUMBER: 47-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2391 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 20.19
Matches 69; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      , MOLECULE TYPE: protein US-08-446-855A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22201-4714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Virginia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                         US-08-446-855A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   249
```

요 ò g õ g

ð

Op

õ

```
91 NKEKR---DKFLSSYNYIKDSIDTDINFANDVLGYYKILSEKYKSDLDSIKKYI--NDKQ 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              146 GENEKYLPFLNNIETLYK---TVNDKIDLFVIHLEAKVLNYTYEKSNVE---IKELIYLK 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5 SVIHNILSKIENEYEVLY----LKPLAGVYRSLKKQLENNVMTFNVNV---KD----IL 52
                                                                                                                                                                                                                                                                                                                                                  CTHER INFORMATION: Description of Artificial Sequence:Polypeptide US-08-973-462-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Kellner, Roland
VENTION: Biosynthetic Process For The Preparation
VENTION: Of Chemical Compounds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 5.8%; Score 114; DB 4; Length 1786; Best Local Similarity 25.4%; Pred. No. 1.2; Matches 67; Conservative 43; Mismatches 84; Indels 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        53 NSP-----FNKRENFKNVLESDLIPYKDLTSSNYVVKDPYK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sterne, Kessler, Goldstein & Fox
                                    PCT/FR96/00894
                              EARLIER APPLICATION NUMBER: PCT/FR96/008
EARLIER FILING DATE: 1996-06-12
EARLIER APPLICATION NUMBER: F 95/07007
EARLIER FILING DATE: 1995-06-13
SOFTWARE: PATENTIN VET: 2.0
SEQ ID NO 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   G tz, Friedrich
Schnell, No. 5837485bert
Augustin, Johannes
Engelke, Germar
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1266 ELEKALSEDSKE----IIDAKDD 1284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 21, Application US/08392625 Patent No. 5837485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              200 TIQDKLADFKKNNNFVGIADLSTD 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1100 New York Avenue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Entian, Karl-Dieter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS
      CURRENT FILING DATE: 1998-02-06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Rosenstein, Ralf
Kaletta, Cortina
Klein, Cora
                                                                                                                                                                                                                                                                 TYPE: PRT ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Wieland, Bernd
Kupke, Thomas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Jung, G nther
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STREET: 1100 New CITY: Washington STATE: D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE:
                                                                                                                                                                                                                                     1786
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-392-625-21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT:
                                                                                                                                                                                                                                                                                                                             FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
Cloning and Expression of Plasmodium faliciparum Transmission-Blocking Target Antigen, Pfs230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 8, Application US/08973462B
Patent No. 6191270
GENERAL NI-FORMATION:
APPLICANT: DRUILHE, PIERRE
APPLICANT: DAUBERSIES, PIERRE
APPLICANT: DAUBERSIES, PIERRE
TITLE OF INVENTION: MALABRIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
FILE REFERENCE: 0660-0125-0 PCT
CURRENT APPLICATION NUMBER: US/08/973,462B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----- AN 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        118 DVLGYYKILSEKYKSDLDSI---KKYIND-----KQGENEKYLPFLNNIETLYKTVND 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        168 KIDLFVIHLEAKVLNYTYEKSNVEIKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHN 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21 LYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRENFKNVLESDLIPYKDLTSSN 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----LFYY----SH 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              83; Indels 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 1; Length 3135;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               :| |::| ||||::|:
-----MKIFSSPKRKKTVVQK 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           228 NLLTKFLSTGMVFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQ 275
                                                                                                                                                                                                                                                                                                                                                                                          CURRENT APPLICATION DATA:

CURRENT APPLICATION DATA:

CURRENT APPLICATION DATA:

APPLICATION UNBER:

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION WHERE:

CLASSIFICATION WHERE:

APPLICATION WHERE:

APPLICATION WHERE:

APPLICATION WHERE:

REGISTANTION WHERE:

TELECOMMUNICATION WHERE:

TELECOMMUNICAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81 YVVKDPYKFLNKEKRDKFLSS---YNYIKD---SIDTDINF---
                           TITLE OF INVENTION: faliciparum Transmission-Blowdbards SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10 LFLIILVYIF-SEKKDLRCNVIKGN-NIKDDEDKRFH-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           41; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 117;
Pred. No. 1
                                                                                                                                                                                                                                                                                                                    MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               : | :: | | 190 DTLIEWFDDNTNEENFLLTFLKRCL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     169 ----IQNVEEKIQRDTYE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
LENGTH: 3135 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 22.2%
                                                                                                                                                                                                 ZIP: 94111-3834
COMPUTER REDABLE FORM:
MEDIUM TYPE: Flore
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: protein
   TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-323-170B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-973-462-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
```

```
SOFTWARE: PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                PRICE APPLICATION DATA:
APPLICATION NUMBER: US 08/392,625
FILING DATE: 22-FEB-1995
FILING DATE: 20-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/876,791
FILING DATE: 30-APR-1992
PRIOR APPLICATION NUMBER: US 07/784,234
FILING DATE: 31-OCT-1991
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Search completed: August 8, 2001, 12:30:01 Job time: 49 sec
                                                                                                                                                                                                 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,961A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
    1100 New York Avenue, NW
                                                                                                                                            COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH: 455 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS:
                                                                                                    COMPUTER READABLE FORM:
                       Washington
                                                             U.S.A.
                                                                                                                           MEDIUM TYPE:
                                                                              20005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-466-961A-21
                                        STATE: C
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                164 --FLNNDALEVTPTNYDIIQGFSGVGRYLLNRISYNYNAKKALKHI--LNYFKTIH--- 215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          207 DFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM---VFENLLKSILSNLLDWKL---ARYV 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           261 KHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQEGSKCVENPNPTCNENNGGCDAD 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       74 KDLTSSNYVVK-----DPYK-----FLNKEKR-----DKFLSSY----NYIK 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              118 SGIDGFSLFSGLSGIGFALD-----IASDKQYSYQSILEQIDNLLVQYVFD----- 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             152 LPFLNN----IETLYKTVNDKIDLFVIHLEAKVLNYTYEKSNVEIKELIYLKTIQDKLA 206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 107 DSID------TDINFANDVLGYYKILSEK---YKSDLDSIK-----KYINDKQGENEKY 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               17 EYEVLYLKPLAGVYRSLKK--QLEN-NVMTFNVNVKDILNSPFNKRENFKNVLESDLIPY 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      216 -YSKDNWLVSNEHQFLDIDKQNFPSGNINLGLAHGILGPLSLTALSKWNGIEIEGHEEFL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Keliner, Roland
TITLE OF INVENTION: Biosynthetic Process for the Preparation of
TITLE OF INVENTION: Chemical Compounds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             51; Mismatches 119; Indels 110;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; DB 2; Length 455; 0.24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sterne, Kessler, Goldstein & Fox
                                                                                                                                                         NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0652.0980002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                    5.7%; Score 113; 22.2%; Pred. No. 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       G tz, Friedrich
Schnell, No. 5843709bert
Augustin, Johannes
Engelke, Germar
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/392,625
                                                         CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/876,791
FILING DATE: 30-APR-1992
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       275 QDFTSFLLKPE-FKNNNEWFDRYD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 21, Application US/08466961A
Patent No. 5843709
GENERAL INFORMATION:
APPLICANT: Entian, Karl-Dieter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Rosenstein, Ralf
Kaletta, Cortina
Klein, Cora
Wieland, Bernd
                                                                                                                                                                                                                                                                                                                                 455 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 22.29
Matches 80; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kupke, Thomas
Jung, G nther
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CORRESPONDENCE ADDRESS: ADDRESSE: Sterne, K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-466-961A-21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                 ;
TOPOLOGY:
US-08-392-625-21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
```

```
19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      164 --FLNNDALEVTPTNYDIIQGFSGVGRYLLNRISYNYNAKKALKHI--LNYFKTIH---- 215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            107 DSID------TDINFANDVLGYXKILSEK---YKSDLDSIK----KYINDKQGENEKY 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   152 LPFLNN----IETLYKTVNDKIDLFVIHLEAKVLNYTYEKSNVEIKELIYLKTIQDKLA 206
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 5.7%; Score 113; DB 2; Length 455;
Best Local Similarity 22.2%; Pred. No. 0.24;
Matches 80; Conservative 51; Mismatches 119; Indels 110; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       74 KDLTSSNYVVK-----DPYK-----FLNKEKR-----DKFLSSY----NYIK 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     207 DFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM---VFENLLKSILSNLLDWKL---ARYV 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           216 -YSKDNWLVSNEHOFLDIDKONFPSGNINLGLAHGILGPLSLTALSKMNGIEIEGHEEFL 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   261 KHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQEGSKCVENPNPTCNENNGGCDAD 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  17 EYEVLYLKPLAGVYRSLKK--QLEN-NVMTFNVNVKDILNSPFNKRENFKNVLESDLIPY 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11 EYSVLILKKKNLFYLFLMKLQKLKNIGMVVININ-------NIKKILENKITFL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ODFTSFLLKPE-FKNNNEWFDRYD----
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/COCKET NUMBER: 0652
TELECOMMUNICATION INFORMATION:
TELEFAN: (202) 371-2600
TELEFAX: (202) 371-2600
INFORMATION FOR SEQ ID No: 21:
```

Lasmodium

```
plasmodium
plasmodium
plasmodium
plasmodium
                                                                                                                                          plasmodium
plasmodium
plasmodium
plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               plasmodium
                                                               plasmodium
                                                                                            plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                          plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-93295445; PubMed-8515786; Jongwutiwes S., Tanabe K., Kanbara H.; Jongwutiwes S., Tanabe K., Kanbara H.; Sequence conservation in the C-terminal part of the precursor to major merozoite surface proteins (MSP1) of Plasmodium falciparum field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
     025727
025728
025728
025719
025719
025719
025720
025720
025721
025721
025722
025722
025722
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
025723
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 99.1%; Score 2033; DB 5; Length 570; Best Local Similarity 99.2%; Pred. No. 2.3e-101; Matches 391; Conservative 0; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  570 AA; 64632 MW; 424BF553CCC2F2BE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 570 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13345; BAA02606.1; -
InterPro; IPR000561; -
Pfam; PF000008; EGF; I.
Merozoite; EGF-like domain.
                                                                                                                                          Q25720
043996
Q25724
Q25724
Q25723
Q25722
Q2572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            003999
025966
025976
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q25984
Q9TYG1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q25971
Q25973
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9TVG8
Q02569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
1930
1925
1916 : 5
1914 : 5
1914 : 5
1916 : 1900
1900
1900
1891
1891
                                                                                                                                                                                                                                                                                                                                       1887.5
1882.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
11254.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 025968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121
     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                     D
                                                                                                                                                                                                                         (without alignments)
743.944 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  025968 plasmodium 091929 plasmodium 025970 plasmodium 025980 plasmodium 025982 plasmodium 025973 plasmodium 025974 plasmodium 025975 plasmodium 025977 plasmodium 025979 plasm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      lasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  the number of results predicted by chance to have a than or equal to the score of the result being printed, by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                       US-09-500-376-3
2052
1 AISVTMDNILSGFENEYDVI......SNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                               ; Search time 70.07 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09tzt4
025983
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q25924
Q25967
                               GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         425026 seqs, 132305027 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Listing first 45 summaries
                                                                                                                                          protein search, using sw model
                                                                                                                                                                                            August 8, 2001, 12:33:42
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sp_unclassified:*
sp_vertebrate:*
sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     091YG2
025970
025980
025980
025923
025978
025978
025977
025977
025977
025977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        025968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_invertebrate:*
sp_mammal:*
sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Post-processing: Minimum Match 0% Maximum Match 100%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_archea:*
sp_bacteria:*
sp_fungi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB seq length: 0
DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_plant:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SPTREMBL_16:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_human:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            569
1720
569
569
569
569
569
1694
1704
569
569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length
                                                                                                                                                                                                                                                                                                                                                                                                                           BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. is the n
score greater than
and is derived by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     999988.
9997...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Total number of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2033
2017.5
2017.5
2017.5
2017.5
2015.2
2011.5
2001.5
2001.5
2001.5
2001.5
2001.5
2001.5
2001.5
2001.5
2001.5
2001.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score
                                                                                                                                                                                                                                                                                                                                       Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                        Scoring table:
                                                                                                                                          OM protein
                                                                                                                                                                                                                                                                                                                                                                    Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Database
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Minimum
                                                                                                                                                                                               Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .
9
                                                                                                                                                                                                                                                                                                          Title:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Result
```

ö

Wed

```
Query Match
                                                                                                                                                025970
                           417
                                                                                   361
                                                                                                   537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 361
                                              301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             301
                           q
                                                                                                                                                          ò
                                                               a
                                                                                  ò
                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-93295445; PubMed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates.";
                                                       300
                                                                                           360
                                                                                                                                                                                                                                                                                                                         malaria parasite
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                         297 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL
                                                       LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                 1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99.1%; Score 2033; DB 5; Length 570; 99.2%; Pred. No. 2.3e-101; tive 0; Mismatches 3; Indels
                                                                                                                                                                                                                                                                    Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8674DEC89B2D662A CRC64;
                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE-88011243; PubMed=3079521;
Tanabe K., Mackay M., Goman M., Scaife J.G.;
"Allelic dimorphism in a surface antigen gene of the
                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                        PROTEIN PRECURSOR (FRAGMENT
                                                                                                                                       PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                               Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13143; BAA02604.1; -.
InterPro; IPR000561; -.
Pfam; PF000008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                                                                                                                                              Created)
                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                 Plasmodium falciparum.";
J. Mol. Biol. 195:273-287(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     64630 MW;
                                                                                                                                                                                                          01-MAY-2000 (TrEMBLrel, 1:
01-MAY-2000 (TrEMBLrel, 1:
01-UUN-2000 (TrEMBLrel, 14
MAJOR MEROZOITE SURFACE PI
MSP1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 99.2
Matches 391, Conservative
                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                          Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      570
                                                                                                                                                                                                                                                                                                                                                                                                                                                             NON_TER
SEQUENCE
                                                                                                                                                                                            Q9TYG2
                   181
                                    357
                                                      241
                                                                        417
                                                                                          301
                                                                                                            477
                                                                                                                              361
                                                                                                                                               537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                357
                                                                                                                                                                                  ò
                                 g
                                                      ò
                                                                    g
                                                                                         ογ
                                                                                                           셤
                                                                                                                       ò
                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              416 LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY 60
241 LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                               HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADATCTEEDSGSSRKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       part of the precursor to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98.3%; Score 2017.5; DB 5; Length 569; 99.0%; Pred. No. 1.5e-100; iive 0; Mismatches 3; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4D761FF472753142 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURPACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                     PDSYPLEDGIFCSSSNFLGISFLLILMLILYSFI 570
                                                                                                                                                                                                                                                                                                                                                                                             Ā
                                                                                                                                                                                                          PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Jongwutiwes S., Tanabe K., Kanbara H.; "Sequence conservation in the C-terminal pamajor merozoite surface proteins (MSP1) of field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mol. Blochem. Parasitol. 59:95-100(1993).
EMBL; D13347; BAA02608.1; -.
Pfam; PF00008; EGF.1.
Merozolte; EGF-11ke domain.
NON_TER
SEQUENCE 569 AA; 64502 MW; 4D761FF472
                                                                                                                                                                                                                                                                                                                                                                                          569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               536 PDSYPLEDGIFCSSSNFLGISFLLILMLILYSFI
                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
MEDLINE=93295445; PubMed=8515786;
                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
Matches 390; Conserv
```

۲;

ή.

Gaps

025980 4

MSP1.

```
Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FLDVLESDLAQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY 60
                                                                                                                                                                                                                                                                                                                                                           1;
                                                                                                                                                                                                                                                                                                                     5; Length 569;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
PRECURSOR OF THE MAJOR MENGOSITE SURFACE ANTIGENS.
Plasmodium falciparum (isolate NF54).
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-NE54;
Tolle R.;
Submitted (JUL-1994) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                   569 AA; 64530 MW; 4D734580516D2142 CRC64;
                                                                                                                                                                                                                                                                                                                     98.3%; Score 2017.5; DB 59.0%; Pred. No. 1.5e-100; iive 0; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PDSYPLEDGIFCSSSNFLGISFLLILMLILYSFI 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 1720 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                         Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13354; BAA02615.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Tolle R., Bujard H., Cooper J.A.; Exp. Parasitol. 0:0-0(1995).
                                                    SEQUENCE FROM N.A.
MEDLINE=93295445; PubMed=8515786;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1069-1720 FROM N.A.
                                                                                                                                                                                                             Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                     Query Match 98.3
Best Local Similarity 99.0
Matches 390; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                            InterPro; IPR000561; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=5843;
                  NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=NF54;
                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                  NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            025922;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          025922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
Q25922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ij
                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
MEDLINE-93295445; PubMed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  237 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     241 LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 569;
                                                                                                                                                                                                               Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8008861DECECD8DC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       025982 PRELIMINARY; PRT; 569 AA. 025982; 01-NOV-19982; 01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-NAY-2000 (TrEMBLrel. 13, Last annotation update) MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT). Plasmodium falciparum.
                                                                                                                     01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                    01, Created)
01, Last sequence update)
13, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98.3%; Score 2017.5; DB 5;
ilarity 99.0%; Pred. No. 1.5e-100;
Conservative 0; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                     569
                                                                                                                                                                                                                                                                                                                                                                          Mol. Blochem. Parasitol. 59:95-100(1993)
EMBL; D13353; BAA02614.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI
                                                                     PRT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64536 MW;
                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                                        (TrEMBLrel.
                                                                                                                                                                                                Plasmodium falciparum,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      569 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 390; Conserv
                                                                                    Q25980;
01-NOV-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
SEQUENCE
```

61

g δ g 121

οŽ

g ò g δ QQ ò g ò g

326

301

536

569;

236

9

240

us-09-500-376-3.rspt

```
241 LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 360
                                                                                                                                                                                                  61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                    121 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEOKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Peterson G.M., Coppel R.L., McIntyre P., Langford C.J., Woodrow G., Brown G.V., Anders R.F., Kemp D.J.;
"Variation in the precursor to the major merozoite surface antigens of Plasmodium falciparum.",
Mol. Biochem. Parasitol. 27:291-302(1988).
                                                                                                                                                177 AISVTMDNILSGFENEYDVIXLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRRY
                                                                                                                                                                                                                 181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                  476 HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK
                                                                                                                             1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
NCBL_TaxID=5833;
                                                       5; Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Tolle R., Bujard H., Cooper J.A.; "Plasmodium falciparum: variations within the C-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Submitted (JUL-1994) to the EMBL/GenBank/DDBJ databases
     B8B9B63EECB1DA51 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
MEROZOITE SURFACE ANTIGEN 1 (FRAGMENT).
                                                     Score 2013.5; DB 5
Pred. No. 2.5e-100;
0; Mismatches 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          536 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  652 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=88142999; PubMed=2449612;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=95354793; PubMed=7628566;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    merozoite surface antigen-1.";
Exp. Parasitol. 81:47-54(1995).
EMBL; 235328; CAA84557.1; -.
InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
     64536 MW;
                                                     98.1%;
98.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 1-298 FROM N.A. STRAIN=FCH5/NF7;
                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Plasmodium falciparum.
     AA;
                                                     Query Match
Best Local Similarity
Matches 389; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A. STRAIN-FCH5/NF7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-FCH5/NF7;
     269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Merozoite
     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ď.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q25923
Q25923;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              rolle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 225923
     g
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                       q
                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                            1507 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNPYDFEAIKK 1566
                           MEDLINE-96123395; PubMed-8577332;
Pan W., Tolle R., Bujard H.;
"A direct and rapid sequencing strategy for the Plasmodium falciparum antigen gene gp190/MSA1.";
Mol. Biochem. Parasitol. 73:241-244(1995).
EMBL: Z35327; CAA84556.1;
Interpro: IPR000561;
Pfam: PF000008; EGF; 1.
Signal; Merozoite.
                                                                                                                                                                                                                                                                                                                                                        major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                             FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121 YYEKVLAKYKODLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                       1 AISVIMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK
                                                                                                                                                                                                                                                                                                    0;
                                                                                                                                                                                                                                                                    Length 1720;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                               195725 MW; 717B2FD1E637A8A3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                       .
9
                                                                                                                                                                                                                                                                 Score 2015; DB 5;
Pred. No. 6.6e-100;
0; Mismatches 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 1720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-93295445; PubMed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13351; BAA02612.1; -.
InterPro; IPR000561; -.
Pen; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                  98.2%;
98.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (TrEMBLrel. 01, (TrEMBLrel. 01,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                 Query Match 98.2
Best Local Similarity 98.5
Matches 388; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Plasmodium falciparum
                                                                                                                                                                                                             SEQUENCE 1720 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1996 (
01-NOV-1996 (
01-MAY-2000 (
                                                                                                                                                                                            SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            025978
```

1387

61

셤 ò q ò 8 ò g õ g

181

241 1567 301 1627 361 1687

g

ò

g

ò

of

region

```
Ą.
                                                                                                                                                                                                                                                                                                                                                                                        536 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 569
                                                                                                                                                                                                                                                                                                                                                                       PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13348; BAA02609.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 98.2
Matches 387; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR000561; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             569 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     025974
                                                                                                                                                                                                                                                          416
                                                                                                                                                                                                                                                                                                                                   476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         297
                                                                                                                                                 181
                                                                          121
                                                                                                                                                                                                                                                                                              301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    025974
                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                8
                                                                        ò
                                                                                                             g
                                                                                                                                             οy
                                                                                                                                                                                g
                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-93295445; PubMed-8515786;
MEDLINE-93295445; PubMed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSPI) of Plasmodium falciparum from field isolates.";
Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13346; BAA02607.1;
InterProc. IPMO0561; --
                                                                                                                                                                                                                                                                                           FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                        498
                                                                                                                                                                                                                                                                                                                                                                                                                                              300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                               9
                                                                                                                                                                319 FLDVLESDLAQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS
                                                                                                                                                                                                                                                                                                                                                                                      1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                   181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                            LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ï
                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Alveolata; Apicomplexa; Haemosporida; Pląsmodium.
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   97.5%; Score 2001.5; DB 5; Length 98.2%; Pred. No. 1.1e-99; Live 1; Mismatches 5; Indels
                                                                                                           Indels
                                                                          Length
                   2B6A87737B490A62 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A26C45914AF01396 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (TrEMBLrel. 01, Created)
(TrEMBLrel. 01, Last sequence update)
(TrEMBLrel. 13, Last annotation update)
OITE SURFACE PROTEIN PRECURSOR (FRAGMENT).
                                                                                                             9
                                                                    Score 2011; DB 5;
Pred. No. 3.9e-100;
1; Mismatches 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                   74292 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               64485 MW;
                                                                        98.0%;
98.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                        Local Similarity 98.2
nes 387; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         387; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Plasmodium falciparum.
                 652 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  025969
025969;
01-NOV-1996 (TrEM
01-NOV-1996 (TrEM
01-NAY-2000 (TrEM
MAJOR MEROZOITE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 269
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NON_TER
SEQUENCE
                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local
                                                                                           Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                177
                                                                                                                                                                                                                                                                                                                                                                                                      439
                                                                                                                                                                                                                                                                                                                                                                                                                                            241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       559
                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
025969
10 025969
10 021
10 021
10 021
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
10 031
 ST
                                                                                                                                                                              g
                                                                                                                                                                                                                                               Ω
                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                             ò
                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
```

```
MEDLINE=93295445; PubMed=8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal part of the precursor to the
major merozoite surface proteins (MSPI) of Plasmodium falciparum from
field isolates.";
                                                                                              HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 YYEKVLAKYKUDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL
                                                                                                                                                                   YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK
                                                                                                                                                                                               LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                              1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 97.5%; Score 2001.5; DB 5; Length 569; 98.2%; Pred. No. 1.1e-99; ive 1; Mismatches 5; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         64457 MW; A2691FE569E80396 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1996 (TrEMBLrel. 01, Created)
U-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURRACE PROFEIN PRECURSOR (FRACMENT)
```

ï

Š

9

025975 025975

1025975 002

476

a

ò a

g ò 536

```
major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    177 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY 236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HIDERERCKCLENYKORGDKCVENPNPTCNPNNGGCDADAKCTEEDSGSNGKKITCECTK 535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    237 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                              Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A26C45914AF01396 CRC64;
                                                                                               01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAY-2000 (TrEMBLrel. 13, Last annotation update)
NAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               97.5%; Score 2001.5; DB 5; 98.2%; Pred. No. 1.1e-99; 1ive 1; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
MEDLINE-91295445; Pubmed-8515786;
Jongwuliwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal parts."
                                                                                                                                                                                                                                                                                                                                                                                               Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13350; BAA02611.1;
InterPro: IPR000561;
Pfam; PF00008; EGF; I.
Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     569
                                                            569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI
                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64485 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 98.2
Matches 387; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       569 AA;
                                                                                                                                                                                                                                        NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   025979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          025979;
                                                          025977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 13
Q25979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MSP1
                  RESULT
Q25977
                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ij
                    MEDLINE-93295445; PubMed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal part of the precursor to the
major merozoite surface proteins (MSP1) of Plasmodium falciparum from
field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                            301 HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121 YYEKVLAKYKODLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLINIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            177 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY 236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     297 YYEKVLAKYKDDLESIKKVIKEEKE-FPSSPPTTPPSPAKTDEOKKESKFLPFLTNIETL 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241 LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                 1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                416 LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 5; Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 97.5%; Score 2001.5; DB 5; Length Best Local Similarity 98.2%; Pred. No. 1.1e-99; Matches 387; Conservative 1; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ED4284B2867C9703 CRC64;
                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TREMBLrel. 01, Last sequence update)
01-NAX-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT).
                                                                                                                                                                                PDSYPLEDGIFCSSSNFLGISFLLILMLILYSFI 569
                                                                                                                                                                                                                                                                                                      Æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                         361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13349; BAA02610.1; -.
InterPro; IPR000561; -.
                                                                                                                                                                                                                                                                                                      269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                569 AA; 64459 MW;
                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                           Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NON_TER
SEQUENCE
```

ò g ò g g

οy

a

ò

1;

Length 569;

5;

Indels

AA

301

g ò q

ò

361

ò

ä

용 οy g δ g ò 셤 ò 8 ò

```
97.5%;
                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                  Best_Local Similarity
Matches 387; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09NHX1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9NHX1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GP195
                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9NHX1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACC OX SEL REL REL DER REL DER REL DER REL DER REL DER REM DER REL DER REM DER REL DER REM DER
                                                                                                                                                                    g
                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                 å
                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1;
                                             SEQUENCE FROM N.A.
MEDLINE-93295445; PubMed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ÷
                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 569;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                Score 2001.5; DB 5; Length
Pred. No. 1.1e-99;
L; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Jiang G., Liu R.Z., Daubenberger C.A., Pluschke G., Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases. EMBL; AF062348; AAC72884.1; -. InterPro; IPR000561; -. Pfam; PF00008; EGF; 1. Merozotte. SEQUENCE 1694 AA; 192794 MW; 84CFCORTOGFFK778 CPACALL.
                                                                                                                                                                                                                                                                                                                                                           A2691FE569E80396 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1; Mismatches
                                                                                                                                                                                              Mol. Blochem. Parasitol. 59:95-100(1993).
EMBL; D1352; BAA02613.1; -.
InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
Merozoite; EGF-11ke domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 1694
                                                                                                                                                                                                                                                                                                                                                           569 AA; 64457 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                97.5%;
98.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches 387; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=5833;
NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-HN1
                                                                                                                                                                                                                                                                                                                                   NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9TZT5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 14
Q9TZT5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MSP1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    476
```

g

ò 셤 

```
240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SENDING.......
STRAIN=FCC1/HN;
Shan Z.X., Yu X.B., Li X.R., Ma C.L., Fang J.M.;
Shan Z.X., Yu X.B., Li X.R., Ma C.L., Fang J.M.;
Shan Z.X., Yu X.B., Li X.R., Ma C.L., Fang J.M.;
"Molecular cloning and sequence analysis of major merozoite surface antigen(gp195)gene of Plasmodium falciparum isolate FCC1/HN.";
Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AF218248; AAF27526.1;
InterPro; IPR000561;
InterPro; IPR000861;
Pfam; PF00008; EGF; 1.
Probom; PD001527; -; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                      1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKOIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                  181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
     Length 1694;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 1694;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
NCBL_TaxID=5833;
                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1694 AA; 192766 MW; B51634A49E0F6728 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-0CT-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
MAJOR MEROZOITE SURFACE ANTIGEN.
     .,
.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 97.5%; Score 2001.5; DB 5; Best Local Similarity 98.2%; Pred. No. 3.4e-99; Matches 387; Conservative 1; Mismatches 5;
Score 2001.5; DB 5;
Pred. No. 3.4e-99;
1; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1661 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 1694
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 1694 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
```

÷

240	
181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK 240	
KIDDYLINLKAKINDCNVEKDEAHV	
181 YNNLVNKIDD	
Qy 18.	

Db 1481 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK 1540
Qy 241 LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300

g

ò

Search completed: August 8, 2001, 12:33:44
Job time: 272 sec

OY 301 HIDERECKCLLNYKQEGDKCRENPNPTCNENNGGCDADARCTEEDSGSKKKITCECTK 360 DD 1601 HIDERECKCLLNYKQEGDKCVENPNPTCNENNGGCDADARCTEEDSGSNGKKITCECTK 1660

```
US-08-290-919-4
                                                                                                     August 8, 2001, 12:29:56; Search time 34.41 Seconds (without alignments) 235.763 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 4, Appli
Sequence 12, Appli
Sequence 2, Appli
Sequence 11, Appli
Sequence 11, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 186, Appli
Sequence 186, Appli
Sequence 186, Appli
Sequence 115, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 58, Appl
Sequence 115, Appl
Sequence 115, Appl
Sequence 120, Appl
Sequence 120, Appl
Sequence 120, Appl
Sequence 6, Appli
Sequence 1, Appli
Sequence 1, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                   1 AISVTMDNILSGFENEYDVI......SNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /cgn2_6/ptodata/2/jaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/jaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/jaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/jaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/jaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/2/jaa/PCTUS_COMB.pep:*
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-290-919-4
US-08-290-919-3
US-08-290-919-12
US-08-290-919-12
US-08-290-919-11
US-08-290-919-11
US-08-257-073-11
US-08-480-604A-6
US-08-75-587-186
US-08-75-587-186
US-08-75-587-186
US-08-75-587-186
US-08-75-587-186
US-08-75-587-187-185
US-08-86-968C-58
US-08-86-968C-58
US-08-86-968C-58
US-08-86-968C-58
US-08-86-968C-58
US-08-86-968C-58
US-08-86-968C-58
US-08-86-968C-193
                                                                                                                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PCT-US95-16216-1
                                                                                                                                                                                                                                                                                                           197339 seqs, 20590346 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                     OM protein - protein search, using sw model
                                                                                                                                                                                                                                                       BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                 Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                         US-09-500-376-3
2052
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Perfect score:
                                                                                                                                                                                                                                                        Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Database :
                                                                                                                                                                                                                     Sequence:
                                                                                                                                                                                                                                                                                                             Searched:
                                                                                                         Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Result
```

Sequence 8, Appli Sequence 8, Appli Sequence 1, Appli Sequence 52, Appl Sequence 51, Appli Sequence 51, Appli Sequence 7, Appli Sequence 2, Appli Sequence 2, Appli Sequence 2, Appli Sequence 6, Appli Sequence 6, Appli Sequence 6, Appli Sequence 6, Appli Sequence 6, Appli Sequence 6, Appli Sequence 5, Appli
US-07-789-915A-8 US-08-005-002C-8 US-08-487-203A-8 US-08-315-793-52 US-09-315-793-52 US-08-592-126-148 US-08-104-324B-4 US-09-104-324B-4 US-09-104-324B-4 US-08-487-8 US-08-568-459A-8 US-08-568-459A-8 US-08-568-459A-8 US-08-290-731C-5 US-08-290-731C-5 US-08-290-731C-5 US-08-290-731C-5 US-08-75-587-184 US-08-75-587-184 US-08-75-587-184 US-08-75-587-184 US-08-75-587-184
111140041100000004
912 912 912 1053 1312 1312 1312 931 944 952 951 157 157 159 699
พพพพพพพพพพพพพพพพพพพพพพพพพพพพพพพพพพพพพ
112.5 112.5 112.5 112.5 111.1 111.1 110.1 110.1 100.5
22 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2

## ALIGNMENTS

```
Sequence 4, Application US/08290919
Patent No. 5720959
Patent No. 5720
```

ö

```
Gaps
    Gaps
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ó
                                           325 PNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPDSYPLFDGIFCSSSN 376
                                                                    2 PNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSN 53
                                                                                                                                                                                                                                       GENERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              281 NISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCEENPNPT 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 NISQHQCVKKQCPQNSGCFRHLDEREECKCLLNYKQEGDKCVENPNPT 48
Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 271; DB 1; I
Pred. No. 7.3e-14;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/6BC8/
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                   E: CUSHMAN DARBY & CUSHMAN, L.L.P. 1100 NEW YORK AVENUE, N.W.
  1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                              Sequence 12, Application US/08290919 Patent No. 5720959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-290-919-2
Sequence 2, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13.2%;
ilarity 95.8%;
Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEO ID NO: 12
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH: 48 amino acids
  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             single
                                                                                                                                                                                                                                                                                                                                                                                                                                        STREET: 1100 NEW CITY: WASHINGTON STATE: D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: amino acid
STRANDEDNESS: si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
Matches 46; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                   ADDRESSEE:
                                                                                                                                                     RESULT 3
US-08-290-919-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-290-919-12
  46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
  Matches
                                             δ
                                                                                   g
                                                                                                                                                                          ó
                                                                                                                                                                            ö
                                                                                                                                                                                                                     325 PNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPDSYPLFDGIFCSSSN 376
                                                                                                                                                                                                                                           2 PNPICNENNGGCDADATCTEEDSGSSRKKITCECTKPDSYPLFDGIFCSSSN 53
                                                                                                                                                                                                                                                                                                                                                                                                          GERERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA TITLE OF INVENTION: VACCINE NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
                                                                                                                              Length 53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 282; DB 1; Length 53; Pred. No. 1.3e-14;
                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
MEDIUW TYPE: FIPOPPY MISK
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
PROOR APPLICATION NUMBER: US/08/290,919
FILING DATE: 22-FEB-1992
PRIOR APPLICATION NUMBER: 12-FEB-1993
ATPONEY/AGENT INPORMATION:
REGISTRATION NUMBER: 16,773
REFERENCE/POCKET NUMBER: 16,773
REGISTRATION NUMBER: 16,773
RELECOMMULICATION SEQ 1D NO: 3:
SEQUIBCE CHARACTERISTICS:
INFORMATION FOR SEQ 1D NO: 3:
SEQUIBCE CHARACTERISTICS:
                                                                                                                            14.5%; Score 298; DB 1; L
100.0%; Pred. No. 8.3e-16;
ive 0; Mismatches 0;
                    /label X . /note = "X = M and N, or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P. STREET: 1100 NEW YORK AVENUE, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /label= X /note= X = M and N,
                                                                                                                                                                                                                                                                                                                                                                   Sequence 3, Application US/08290919
Patent No. 5720959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13.7%;
94.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME/KEY: Modified-site
                                                                                                                     Ouery Match
Best Local Similarity 100.0
Matches 52; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            53 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; LOCATION: 1
; OTHER INFORMATION:
; OTHER INFORMATION:
US-08-290-919-3
               ) OTHER INFORMATION:
) OTHER INFORMATION:
US-08-290-919-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CITY: WASHINGTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STATE: D
COUNTRY:
                                                                                                                                                                                                                                                                                                                                 RESULT 2
US-08-290-919-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH:
                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                             g
```

ô

```
12.9%; Score 265; DB 1; Length 48; 95.7%; Pred. No. 2e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   282 ISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCEENPNPT 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 ISQHQCVKKQCPQNSGCFRHLDEREECKCLLNYKQEGDKCVENPNPT 48
                                                                                                                                                                                     CURRENT APPLICATION DATA:

CURRENT APPLICATION DATA:

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 05/08/290,919

FILING DATE: 04-0CT-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/0821.5

FILING DATE: 22-FEB-1992

PRIOR APPLICATION NUMBER: 02-FEB-1993

ATONNEY/AGENT INPORMATION:

FILING DATE: 22-FEB-1993

ATONNEY/AGENT INPORMATION:

NAME: CONULLS, PAUL N.

REGISTRATION NUMBER: 16,773

REGISTRATION NUMBER: 16,773

REGISTRATION NUMBER: 16,773

FELEPHONE: (202) 861-3000

TELEFRANCE/COCKET NUMBER: 16,773

FELEPHONE: (202) 861-3000

TELEFRAN: (202) 861-3000

TELEFRAN: (202) 822-0944

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OTHER INFORMATION: /label= X
OTHER INFORMATION: /note= "X = M and N, or N"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1; Mismatches
                                                                                                                                                  COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 11, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY: Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 95.79
Matches 45; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH: 48 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE: protein
                                                                                                   COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: amino acid
STRANDEDNESS: sir
TOPOLOGY: linear
         WASHINGTON
                                                         20005-3918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-290-919-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LOCATION:
                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-290-919-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
0
                           CHAPPEL, JONATHAN A. FELATING TO MALARIA FENTION: IMPROVEMENTS IN OR RELATING TO MALARIA FENTION: VACCINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
NUMBER OF SEQUENCES: 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         282 ISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCEENPNPT 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2 ISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCVENPNPT 48
                                                                                                                                                                                                                   STAIL:
COUNTRY: U.S.A.
2IP: 2005-3918

ZIP: COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
OPERATING SYSTEM: PC-DOS/MS-DOS
OPERATING PATENTIN Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION NUMBER: PCT/GB93/00367
FILING APPLICATION NUMBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
RECISTATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 16,773
FELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEFAX: (202) 861-3000
TELEFAX: (202) 822-0944
TELEFAX: (202) 822-0944
INFORMATION FOR SCO ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TEPE: amino acids
TEPE: amino acids
STRANDENESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 268; DB 1;
Pred. No. 1.2e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OTHER INFORMATION: /label= X
OTHER INFORMATION: /note= "X = M and N, or US-08-290-919-2
                                  TITLE OF INVENTION: IMPROVEMENTS IN OR RELATITILE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 19
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
STREET: MASHINGTON
STATE: D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     E: CUSHMAN DARBY & CUSHMAN, L.L.P. 1100 NEW YORK AVENUE, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                         CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-290-919-1
Sequence 1, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
BLACKMAN, MICHAEL J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 13.1%;
Best Local Similarity 97.9%;
Matches 46; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME/KEY: Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

```
Gaps
ö
                                                                                                                                                                                                                                                                                               APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDLUM TYPE: Floppy disk
COMPUTER: IBM PC COMPUTER:
OPERATING SYSTEM: PC-DOS/MS-DO
```

ö

```
NAME: Frommer, William S.
REGISTATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2570
TELECOMMUNICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       263 FPNTIISKLIEGKFQDMLNISQHQCVKKQ---
                                         US 07/672,183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 6, Application US/08480604A Patent No. 5736139 GENERAL INFORMATION:
                                                                                                                                                                         TELEPHONE: (212) 840-333
TELEFAX: (212) 840-0712
TELEX: 42506 CURTMS
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 740 amino acids
  18-MAR-1992
                                                        20-MAR-1991
                                                                         ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     308 CKCLLNYKQEGDKCEE 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             680 ЕКЕКЕКЕКЕКЕКЕЕК 695
                    PRIOR APPLICATION DATA:
APPLICATION NUMBER: (
FILING DATE: 20-MAR-1
                                                                                                                                                                                                                                                                                                                            single
                                                                                                                                                                                                                                                                                                       TYPE: amino acid
STRANDEDNESS: Sin
                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   87 KLLNSEQKNI---
                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE:
FRAGMENT TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            204 V-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-480-604A-6
                                                                                                                                                                                                                                                                                                                                                                                                            US-08-257-073-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         278 DMLNIS-QHQCV-KKQCPENSGCFRHLDEREECKCLLNYKQ-EGDKCEENPNPTCNENNG 334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 106;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   32; Indels
                                                                                                                                                                                                                                      REFERENCE/DOCKET NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 21242/HCM/MJL/6BC8/
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (5102) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENOTH: 106 amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     335 GCDADATCTEEDSGSSRKKITCECTKPDSYPLFDGIFCSSSN 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                65 GCDPTASCQNAESTENSKKIICTCKEPTPNAYYEGVFCSSSS 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OPERATING SYSTEM: PC_DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12.5%; Score 255.5; DB 1
47.1%; Pred. No. 2.8e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           E: Curtis, Morris & Safford, P.C. 530 Fifth Avenue, 25th Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 48; Conservative 19; Mismatches
                                                                                            PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
                CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY: UNITED STATES OF AMERICA 2IP: 10036
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/075,783
FILING DATE: 11-UN-1993
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: US 07/852,305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US/08/257,073
FILING DATE: 09-JUN-1994
CLASSIFICATION: 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 5, Application US/08257073
Patent No. 5766597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS ANG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Paoletti, Enzo
APPLICANT: de Taisne, Charles
APPLICANT: Tine, John A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STATE: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-290-919-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-257-073-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ô
```

```
517 VVPPTQSKKKNKNETVSGMDENFDNHPENYFKEEYYYDENDDMEVKVKKIGVTLK---K 572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       226 IDLFKNTNDFEAIK-----STGLVQI 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           278 NATAYELGIVPVNDVLNVNMKNMITGNNFMDVVKSTLAQSGGLGSNDLINFLNQGKEIGE 337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        152 PTTPPSPAKTDEQKKESKFLPFL-----TNIETLYNNLVNKIDDYLINLKAKINDCN 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---EKDEAHVKITKLS-DLKAIDDK 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----CPENSGCFRHLDEREE 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      633 FQQNILE------NDVLNQETEEEMEKQVEAITKQIEAEVDALAPKNK-----EEEEK 679
                                                                                                                                                                                                                                                                                                                   338 NLENITKMNLGDKNNLESFPLDQLNMLKDNLINYEFILNNLKTSVLNKLKDLLLRLLYKA 397
                                                                                                                                                                                                                                                                                                                                                                                118 GISY----YEKVL------AKYKDDLE------SIKKVIKEEKEK-FPSSP 151
                                                                                                                              39 NIITFNLNL----NDILNSRLKKR---KYFLDVLESDLMQFKHISSNEYI-----IEDSF 86
   Length 740;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
TAPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
                                                                Indels
                                                                                                                                                                                                                                                         ---LLK----SYKYI----KESVENDIK-
Query Match 6.2%; Score 127.5; DB 1;
Best Local Similarity 21.8%; Pred. No. 0.1;
Matches 95; Conservative 58; Mismatches 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBER OF SEQUENCES:
```

```
-----VLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPA----- 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    85 SF-KLLNSEQKNI--LLKSYKYIK-----ESVENDIKFAQEGIS-----YYEK- 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  93; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    36 IEKNIITFNLNLNDI-LNSRLKKRKYFL------DVLESDLMQFKHISSNEYIIED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 2710;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                      GENERAL INFORMATION:
APPLICANT: WILLIAMS, JAMES A.
APPLICANT: WILLIAMS, JAMES A.
TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
TITLE OF INVENTION: NEUROTOXIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SOFTWARE: Patentin Release #10, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405,496A FILING DATE: US/08/405,496A FILING DATE: US/08/405,496A FILING DATE: US/08/405,496A FILING DATE: US/08/405,496A PROOR APPLICATION NUMBER: US 08/329,154 FILING DATE: 25-OCT-1994 PRIOR APPLICATION NUMBER: US 08/161,907 FILING DATE: 02-DEC-1993 PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321 FILING DATE: 04-DEC-1992 PRIOR APPLICATION NUMBER: US 07/429,791 FILING DATE: 31-OCT-1989 ATPONENTY OF THE US OF THE US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 6.2%; Score 127.5; DB 2; Best Local Similarity 23.4%; Pred. No. 0.52; Matches 74; Conservative 47; Mismatches 102;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3: MEDLEN & CARROLL, LLP
220 MONTGOMERY STREET, SUITE 2200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01308
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                           Sequence 6, Application US/08405496A
Patent No. 5919665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (415) 705-8410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : 2710 amino acids amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEFAX: (415) 397-8338 INFORMATION FOR SEQ ID NO:
                                                                                                        1008 QLVNLISNAVNDTINV 1023
                                         267 IISKLIEGKFQDMLNI 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CITY: SAN FRANCISCO
STATE: CALIFORNIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; MOLECULE TYPE:
US-08-405-496A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94104
                                                                                                                                                                                                         RESULT 9
US:08-405-496A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        125
                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               956 S-QVNTLNAAFFIQSLIDYSSN-----KDVLNDLSTSVKVQLYAQLFSTGLNTIYDSI 1007
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLIND---DTKKDMLGKLLSTGLVQIFPNT 266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   85 SF-KLINSEQKNI--LLKSYKYIK----ESVENDIKFAQEGIS------YYEK- 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------VLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPA----- 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IEKNIITFNLINLINDI-LNSRLKKRKYFL------DVLESDLMQFKHISSNEYIIED 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -KTD-----EQKKE--SKFLPFLT-NIETLYNNLVNKIDDYLINLKAKINDCNVEKDEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 1; Length 2710;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    47; Mismatches 102; Indels
                                                                                                                                                                                                                                                                                                                                                       CURRENT APPLICATION DATA:

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
FILING PAPLICATION NUMBER: US/08/422,711
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION NUMBER: US 07/985,321
FILING DATE: 31-OCT-1989
ATTONING APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTONING APPLICATION UNMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTONING APPLICATION: UNMBER: US 07/429,791
                        SEE: MEDLEN & CARROLL, LLP: 220 MONTGOMERY STREET, SUITE 2200 SAN FRANCISCO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6.2%; Score 127.5; Di 23.4%; Pred. No. 0.52;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: 0PHD-01763
TELECOMMUNICATION:
TELEPHONE: (415) 705-8410
                                                                                                                                                                      UNITED STATES OF AMERICA
                                                                                                                                                                                                                       COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 6:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2710 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 23.4%
Matches 74; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        protein
CORRESPONDENCE ADDRESS:
                                                                                                    COUNTRY: UNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              160
```

DP OY OD 9 % Q

ò

```
LL--STGLVQIF---PNTIISKLI--EGKFQDMLNISQHQCVK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                E: Bell Seltzer Park & Gibson
310 UCB Plaza, 3605 Glenwood Avenue,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 952555.0
FILING DATE: 14-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9617961.9
FILING DATE: 28-AUG-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5405-135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRILING APPLICATION DATA:
PRICA APPLICATION NUMBER: GB 9523959.6
FILING DATE: 23-NOV-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: US/08/755,587
FILING DATE: 25-NOV-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Materials and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 186, Application US/08755587 Patent No. 6045997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Ashworth, Alan
APPLICANT: Stratton, Michael R
TITLE OF INVENTION: Materials and
TITLE OF INVENTION: identificatio
TITLE OF INVENTION: susceptibilit
NUMBER OF SEQUENCES: 222
CORRESSERE: Bell Seltzer Park &
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Phillip A.
Richard F
                                                                                                                                                                                                                                                                                                                                                                              --CNENNGG---C-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             359 TKPD-SYPLFDGIFCSSSNF 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1494 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Kenneth D Sibley
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STATE: NC
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: Futreal, APPLICANT: Wooster,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER: IBM PC
OPERATING SYSTEM:
                             ---DYLINL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Raleigh
                                                                                                                                                                                                                                                                                                                                                                              324 NPNPT----
                                                                                                              206 KDE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-755-587-186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-755-587-186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STREET:
CITY: Ra
                           188 ID--
                                                                                                                                                                                                    254
                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       32;
                                                                                                                                                                SE S-QVNTLNAAFFIQSLIDYSSN-----KDVLNDLSTSVKVQLYAQLFSTGLNTIYDSI 1007
                                                                      160 -KTD-----EQKKE--SKFLPFLT-NIETLYNNLVNKIDDYLINLKAKINDCNVEKDEA 209
                                                                                                                                210 HVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLIND---DTKKDMLGKLLSTGLVQIFPNT 266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        844 EPVKNIIHNSIDDLIDEFNLLENVSDELYELKKLNNLDEKYLISFEDISKNNSTYSVRFI 903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       134 KFKR--AEDYPIDLYYLMD-LSY----SMKDDLENVKSLGTDLMNEMRRITSDFRIGF 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  185 GSFVEKTVMPYISTTPAKLRNPCTSEQNCTTPFSYKNVLSLTNKGEVFNELVGKQRISGN 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       --PPTTPPSPAK----TDEQKKESKF----LPFLTNIETLYNNLVNK---- 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             48 NDILNSR----LKKRKYFLDVLESDLMQFK-----HISSNEYIIEDSFKLLNSEQKNILL 98
                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: Sheppard, Dean
APPLICANT: Quaranta, Vito
APPLICANT: Quaranta, Vito
APPLICANT: Quaranta, Vito
TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels 189;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99 KSYKYIKESVENDIKFAQEGISYYEKVLAKYKDDLESIKKV---IKEEKEKFPSS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 798;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 4370 La Jolla Village Drive, Suite 700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURREW APPLICATION DATA:
APPLICATION NUMBER: US/O7/728,215
FILING DATE: 19910711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 6.2%; Score 126.5; DB 2; Best Local Similarity 21.0%; Pred. No. 0.13; Matches 105; Conservative 69; Mismatches 137;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATTORNEY AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P31 8717
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEPHONE: (619) 535-8949
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    United States of America
                                                                                                                                                                                                                                                                                                                                                                              Sequence 30, Application US/07728215 Patent No. 5962643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH: 798 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
                                                                                                                                                                                                                                                      1008 QLVNLISNAVNDTINV 1023
                                                                                                                                                                                                                      267 IISKLIEGKFQDMLNI 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CITY: San Diego
STATE: Californi
                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY: UN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
US-07-728-215-30
                                                                                                                                                                                                                                                                                                                                                       US-07-728-215-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    151 ----
  용
                                                                                   g
                                                                                                                                                                         셤
                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                             ð
                                                                                                                                   ð
```

```
BRCA2 cancer
--ND--CNVE 205
                                                                                                                                                                                                : : : | | : : | : : | : : 359 NNMYTMSHYYDYPSIAHL-VQKLSE----NNIQTIFAVTEEFQPVYKELKNLIPKSAVGT 359
                                                                                                                                                                                                                                                                                                                                                        360 LSANSSNVIQLIIDAYNSLSSEVILENGKLSEGVTISYKSYCKNGVNGTGENGRKCSNIS 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    420 IGDEVQFEISITSNKCPKKDSDSFKIRPLGFTEEVEVILQYICECECQSEGIPESPKCHE 479
                                                                        245 LDSPEGGFDAIMQVAVCGSLIGWRNVTRLLVFSTDAGFHFAGDGKLGGIVLPNDGQCHLE 304
                                                                                                                                                AHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDDTKKDMLGK 253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                290 -----ECKCLLNYKQCP-ENSGCFR-----HLDERE-----ECKCLLNYKQEGDKCEE 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 480 G-NGTFECGACRCNEGRVGRHCECSTDEVNSEDMDAYCRKENSSEICSNNGECVCGQCVC 538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PO Drawer 31107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     identification and sequencing of the susceptibility gene and uses thereof 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PatentIn Release #1.0, Version #1.25 (EPO)
```

```
1135
                                                                                                                                                                                                                                                                                                         294
                                                 149
                                                                                                                                  205
                                                                                                                                                                                                                                                                                                                                                                                            324
                                               ŏ
                                                                                       qq
                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ο̈́
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        엄
                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                            ò
                                               21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FERUERAL INFORMATION:

APPLICANT: AUDONNET, Jean-Christophe
APPLICANT: BAUDU, Philippe
TITLE OF INVENTION: PARAINED
TITLE OF INVENTION: PARAINED
TITLE OF INVENTION: PARAINELUBNIA 2 VIRUS
TITLE OF ILUNG PATE: 1991-12-16
CURRENT FILLING DATE: 1991-12-16
EARLIER APPLICATION NUMBER: 960842
EARLIER PILLING DATE: 1996-06-27
EARLIER PILLING DATE: 1997-06-23
NUMBER OF SEQ ID NOS: 31

SOFTWARE: PATEONIN VET: 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N
                                                                                                                                                                                                                                                                                                                                                                                                                                                           |: | :: ||: : ||: || SDNLYKQTENLKISDHA----SQKVDVHENTE------KETAKKPTMYTNQSTYSA 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 260 VQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGD 319
                                                                                                                                                                                                                                                                                                                                             152 PTTPPSPAKTDEQKKESK------FLPFLTNIETLYNNL-VNKIDDYLINLKAKIN 200
                                                                                                                                                                                                                                                                                                                                                                                     921 TNFSHRGAKMSKDREECKDGRELACGTTEITTTPEYEETHSSLEKKKVSNEIAALRPRLL 980
                                                                                                                                                                                                                                                                                                                                                                                                                                  201 DCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDDTKKDMLGKLLST-GL 259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   920 TITKKONLSNPLIELWFKELSTYNKTNENVESLKTDISKNILLFSTKNNSDNFYNDFLLG 979
                                                                                                                                                                                                                 812 WTEEELNNFSDSLNSELLPGIDIKKTDISNHEVIE----NTERKDKITKESLIGTENI 865
                                                                                                                                                                                                                                                            LLKSYKYIKESVENDIKFAQEG - - ISYYE - - - KVLAKYKDDLESIKKVIKEEKERFPSSP 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           866 DTTITKPISTEANNLKSVSTSIKPPKNLKKKLLKSKCRDNVIY-----RPYFSQLEINC 919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----IKFAQEGISYYE--KVLAK------YKDDLESIKKVIKEEKEFF 148
                                                                                                                           760 VIMSNKQQLTANTGQNIKDFDTFYL-----SFQTASRKNIRVSRESLNKARSLLNQK 811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----LESDLMQ--FKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVEND---- 111
                                                 Gaps
                                                                                     4 VTMDNILSGFEN-----EYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLN---DILNSR 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18 DVIYLKPLAGVYRSLK------KQIEKNIITFNLNLNDILNSRLKKRKYF--LDV-- 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1081 DYVENPSCGNSSNSAITENDKNHLSEKQGST---YLSNSTMSNSYSYHPG-FCHSS 1132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           320 KCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPDSYPLFDGIFCSSS 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   65; Mismatches 153; Indels 212;
    Length 1494;
                                                                                                                                                                        55 LKKRKY--FLDVLESDLMQFKHIS----SNEYIIEDSFKLLNSEQK----
5.9%; Score 120.5; DB 3; 22.6%; Pred. No. 0.81; ive 68; Mismatches 167;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. 1.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 5.8%; Score 118;
Best Local Similarity 19.2%; Pred. No. 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 2, Application US/09213053
Patent No. 6159477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; ORGANISM: Canine herpesvirus US-09-213-053-2
                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 102; Conservative
                      Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 1394
                                        94;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-213-053-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQ ID NO 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: PRT
                    Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           112
                                                                                                                                                                                                                                                            97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                 Db
                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                        ò
                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
```

```
APPLICANT: REMY, Elisabeth
APPLICANT: RITZENTHALER, Paul
TITLE OF INVENTION: Nucleic acid sequence and plasmids
TITLE OF INVENTION: comprising at least one phage resistance mechanism,
TITLE OF INVENTION: bacteria in which they are present, and their use
980 IQNQPVNYKLYGSQFYDNGNILLNIKSVDFKTSGIYTWKLYNSNNESIFETFKIQVYAYH 1039
                                                                                                                                                                                                          -----FSNTVIPKVYYKPFISIKFDNGLAICDAKCVSRENNKLQWLVKDIP 1180
                                                                                                                                                                                                                                                                                                                                                                                                                                  1181 INGDDIISGPCLNHPGLVNIQNKIDISDYDEPVTYKCSIIGYPIIFPNFYDEKVFDASDE 1240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----LDEREECKCLL-----NYKQEG--DKCEE 323
                                                                                                                                                                               EKDEAH------VKITKLSDLKAIDD-----KIDLFKNTNDFEAIKKLINDDTKK 248
                                                                                                                                                                                                                                                                                249 DMLGKLLSTGLVQIFPNTIISKL----IEGKFQDMLNISQHQCVK------KQCP 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NPNPT-----DADATC 342
                                                                           SSPPTTPPSPAKTDEQKKE----SKFLPFLTNIETLYNNLVNKIDDYLINLKAKINDCNV 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----LCVTLGFFAAGILLTIAAIIFTFI 1340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              343 TEEDSGSSRKKITCECTKPDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                              1040 SPNVNLKSNPSLYNENYSAICTIANYFP-LESTEIFWFNDGQPIDKKYI----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STREET: 625 Slaters Lane - Fourth Floor CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: US/08/286,325A FILING DATE: 04-AUG-1994 CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: FICHTER, Richard E
REGISTRATION NUMBER: 26,382
REFERENCE/DOCKET NUMBER: REF/BEDL/
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; Sequence 2, Application US/08286325A
; Patent No. 5658770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION WUMBER:
FILING DATE: 09-AUG-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                       EN-----SG-CFRH-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BACON & THOMAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PREVOTS, Fabien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX: (703) 683-1080 INFORMATION FOR SEQ ID NO: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    : 540 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TITLE OF INVENTION: NUCL
TITLE OF INVENTION: COME
TITLE OF INVENTION: back
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         || : ::|| ||:|
|1301 DEETTNLNQKKSTCQC---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; MOLECULE TYPE: protein US-08-286-325A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE: Virginia
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
APPLICANT: PREVOT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-286-325A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGIH:
```

75;

```
EYI -- IEDSFKLLNSEQ-KNILLKSYKY----IKESVENDIKFAQEGISYYEKVLAKYKD 131
                                                                                                                                                                                                                                                                                                                                                                                       309 FANMIRDKVKREKNRILTT------KCEDIIRYIKLFS-----KNRINDENNKVEEV 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        252 GKLLSTGLVQ----IFPNTIISKLIEGKF-QDMLNISQHQCVK--KQCPENSGCFRHLD 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 398 SKLLALRAVKNWKCYSLTNVSMYKKIKGVIVMDMVDYISTNILKYHKQLYDKMSTFEYKR 457
                                                                                                                                                                                                                                  132 DLESIKKVIKEEKEKFPSSPPTTPPSPÅKTDEQKKESKFLPFLTNIETLYNNLVNKIDDY 191
                                                                                                                                                                                                                                                                                                                                                                                                                                           192 LINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDDTKKDML 251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               355 LIHID-----NVSKNNK----LSLSDISSLMDQFRL----NPCTIRNILLSSATIK--- 397
                                                                                                                                                                                                     22 LKPLAGUYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVLE---SDLMQFKHISSN 78
                                                                                                                          Length 677;
                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,640A
FILING DATE: UN-JUN-1995
CLASSIFICATION: 424
                                                                                                                      5.6%; Score 115.5; DB 3;
21.6%; Pred. No. 0.69;
tive 59; Mismatches 123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 193, Application US/08480640A
Patent No. 6033904
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David E.
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE: John P. White STREET: 1185 Avenue of the Americas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----LLNYK 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :: ||| :||| 485 BIKSCKCSICSDSITHHIYETTSCINYK 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME: White, John P
REGISTRATION NUMBER: 28,678
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ. ID NO: 193
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATTORNEY/AGENT INFORMATION:
                                                                                                                        Query Match 5.6%
Best Local Similarity 21.6%
Matches 71; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy di
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; MOLECULE TYPE: protein US-08-480-640A-193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino acid
POSITION IN GENOME:
MAP POSITION: ~2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       304 EREECKC----
                                                            US-08-480-640A-115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-480-640A-193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH:
                                                                                                                                                                                                                                                                                  79
                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                              18;
                                                                                                                                                                                                                          SFKLLNSEQKNILLKSYKYIKESVEN-DIKF--AQEGISYYEKVLAKYKDDLESIK--KV 139
                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                              140 IKEE-----KEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYN--NLVNKIDDY 191
                                                                                                                                                                                                                                                                                                                                                            192 LINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDDTKKDML 251
                                                                                                                  VYRSLKKQIEKN---IITFNLNLNDILNSRLKKRKYFLDVLESDLMQFKHISSNEYIIED 84
                                   Similarity 25.0%; Pred. No. 0.48;
0; Conservative 45; Mismatches 89; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,640A
FILING DATE: US-JUN-1995
CLASSIFICATION: 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --SKLIEGK 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 115, Application US/08480640A
Patent No. 6033904
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David E.
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          E: John P. White
1185 Avenue of the Americas
                                                                                                                                                                                                                                                                                                                                                                                                                                       252 GKLLSTGLV-----QIFPNTII-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: White, John P
REGISTRATION NUMBER: 28,678
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-05.25
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:
LENGTH: 677 amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INDIVIDUAL ISOLATE: S-SPV-001
IMMEDIATE SOURCE:
CLONE: 515-85.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA (genomic)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM: Swinepox virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRANDEDNESS: double
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRANGE 1 linear TOPOLOGY: linear MOLECULE TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New York
: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE: 07
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Kasza
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-480-640A-115
                                     Query Match
Best Local Simi
Matches 70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STATE: Ne COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CITY:
                                                                                                                                                                                                 85
                                                                                                                                                                                                                                                                                                             Q
                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
16;
                                                                                                                                     252 GKLLSTGLVQ-----IFPNTIISKLIEGKF-QDMLNISQHQCVK--KQCPENSGCFRHLD 303
                                                                                                                                                                                                                                                                                                                                                                      398 SKLLALRAVKNWRCYSLTNVSMYKKIKGVIVMDMVDYISTNILKYHKQLYDKMSTFEYKR 457
                                                                                                                                                                                                         132 DLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLUNKIDDY 191
                                                                                                                                                                                                                                309 FANMIRDKVKREKNRILTT------KCEDIIRYIKLFS-----KNRINDENNKVEEV 354
                                                                                                                                                                                                                                                                             192 LINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDDTKKDML 251
                                                                                                                                                                                                                                                                                                    22 LKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVLE---SDLMQFKHISSN 78
Query Match 5.6%; Score 115.5; DB 3; Length 677; Best Local Similarity 21.6%; Pred. No. 0.69; Matches 71; Conservative 59; Mismatches 123; Indels 75;
                                                                                                                                                                                                                                                                                                                                                                                                                    304 EREECKC------LLNYK 315
                                                                                                                                                                                                                                                                                                                                                                                                                                         :: ||| : ||| 485 BIKSCKCSICSDSITHHIYETTSCINYK 485
                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                   g
                                                                                                                                       ŏ
                                                                                                                                                                QQ
                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                    δ
```

Search completed: August 8, 2001, 12:29:58 Job time: 46 sec

```
; Search time 60.73 Seconds
(without alignments)
393.312 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                   1 AISVTMDNILSGFENEYDVI.....SNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*/SIDSB/gcgdata/geneseq/geneseqp/AA1993.DAT:*/SIDSB/gcgdata/geneseq/geneseqp/AA1994.DAT:*/SIDSB/gcgdata/geneseq/geneseqp/AA1994.DAT:*/SIDSB/gcgdata/geneseq/geneseqp/AA1995.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
/SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
/SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /gcgdata/geneseq/geneseqp/AA2000.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .DAT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             . DAT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         412676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | SIDS#gcgdata/geneseq/geneseqp/AA1980.DA7
| SIDS#gcgdata/geneseq/geneseqp/AA1981.DA7
| SIDS#gcgdata/geneseq/geneseqp/AA1982.DA7
| SIDS#gcgdata/geneseq/geneseqp/AA1983.DA7
| SIDS#gcgdata/geneseq/geneseqp/AA1981.DA7
| SIDS#gcgdata/geneseq/geneseqp/AA1985.DA7
| SIDS#gcgdata/geneseq/geneseqp/AA1985.DA7
| SIDS#gcgdata/geneseq/geneseqp/AA1985.DA7
| SIDS#gcgdata/geneseq/geneseqp/AA1989.DA7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /SIDS8/gcgdata/geneseq/geneseqp/AA1989.
/SIDS8/gcgdata/geneseq/geneseqp/AA1990.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /SIDS8/gcgdata/geneseq/geneseqp/AA199
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           412676 seqs, 60623988 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Maximum Match 100%
Listing first 45 summaries
                                                                                                          OM protein - protein search, using sw model
                                                                                                                                                                 August 8, 2001, 12:31:53
                                                                                                                                                                                                                                                                                                                                                                                           BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A_Geneseq_0601:*
                                                                                                                                                                                                                                                                              US-09-500-376-3
2052
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIDS8
                                                                                                                                                                                                                                                                              Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                              Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Searched:
                                                                                                                                                                 Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Joseph Line	Ct the ctoll	P. falciparum synt	derozoite surface	Merozoite surface	rozoite surface	rozoite surface	dified merozoite	dified merozoite	uence of the P1	rozoite surface	MSP1(p19)A prote	PfMSP1(p19)A prote
	800	200	Ч	Me	. We	Me	Me	W	W	Sec	Me	Pf	Jd
	£		AAW54145	AAY 09372	AAY05832	AAY09373	AAY05833	AAY09374	AAY05834	AAP50777	AAB37610	AAW36103	AAW22592
	H.C.	;	19	20	20	20	20	20	20	9	22	18	18
	Query Match Length DR ID		1639	355	355	361	361	376	376	1654	262	116	116
ф	Query		8.09	56.4	56.4	56.4	56.4	56.0	56.0	53.2	31.3	30.0	30.0
	Score		1248.5	1156.5	1156.5	1156.5	1156.5	1149.5	1149.5	1091	641.5	615	615
	Result		7	7	m	4	'n	9	7	æ	σ	10	11

Merozoite surface Merozoite surface PEMSP1(p19)S prote PEMSP1(p19)S prote PEMSP1(p19)S prote MSP1EGF2B EGF2-lik MSP1EGF2B EGF2-lik MSP1EGF1B EGF1-lik Plasmodium falcipa Plasmodium falc
4
AAB37609 AAB37608 AAW32593 AAR41356 AAR41355 AAR41355 AAR41355 AAR41355 AAR118294 AAR118294 AAR118294 AAR118294 AAR118294 AAR118236 AAR18236 AAR18256 AAR18256 AAR18256 AAR18256 AAR18256 AAR18256 AAR18250 AAR18250 AAR18250 AAR18250 AAR18250 AAR18250 AAR1823 AAR18272
2238884444424111111111111111111111111111
108 96 1127 1127 127 127 127 127 137 137 137 137 137 137 137 137 137 13
22222111111111111111111111111111111111
528 528 528 521 304 288 273 273 273 274 275 271 275 134 132 134 132 132 123 123 124 127 127 127 127 127 127 127 127 127 127
11111111111111111111111111111111111111

7

## ALIGNMENTS

gpl90; malaria; MSP-1; merozoite surface protein; stability; vaccine; monoclonal antibody; passive immunisation; parasite. Recombinant production of complete gp190/MSP-1 Plasmodium surface protein - useful in anti-malaria vaccines, also stabilising genes by reducing their AT content P. falciparum synthetic gp190 protein. AAW54145 standard; Protein; 1639 AA. 97WO-EP05441. 96DE-4040817. Bujard H, Pan W, Tolle R; N-PSDB; AAV21451, AAV35363. (first entry) Plasmodium falciparum WPI; 1998-240088/21 (BUJA/) BUJARD H. WO9814583-A2 02-OCT-1997; 02-OCT-1996; 23-SEP-1998 09-APR-1998 Synthetic. AAW54145; AAW54145 RESULT 

```
The present sequence represents a 42 kDa C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against plasmodium falciptum. The nucleic acid sequence encoding MSP-1-42 has been modified (see AAX56008) compared to the native sequence (see AAX56009) such that 306 nucleotide positions have been replaced to lower the AT content (from 76 to 49.7%) and to eliminate 10 mRNA instability motifs while maintaining the same protein amino acid sequence. These alterations allow MSP-1-142 to be expressed in mammalian cell culture and in transgenic mice. Native MSP-1-12 is known to be difficult to express in cell culture systems.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            mammalian cell culture systems or in transgenic animals. The invention allows expression of ,MSP-1 protein in the milk of transgenic animals, and also provides a DNA vaccine comprising
                                                                                                                                                                                                                       Modified malarial protein for use in anti-malarial vaccines
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               rector containing the altered MSP-1-42 sequence
                       (GENZ ) GENZYME TRANSGENICS CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                355
                                                                                                                                                                                                                                                                               Example; Fig 1; 35pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY05832 standard; Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 56.4
Best Local Similarity 59.0
Matches 219; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SYPLFDGIFCS 373
                                                                                                                                    WPI; 1999-288313/24
                                                                               Meade H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ¥,
                                                                                                                                                                  P-PSDB; AAX56008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                02-AUG-1999
                                                                               Chen LH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY05832;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY05832
ID AAY0
XX
AC AAY0
XX
DT 02-A
XX
DE Mero
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δy
                                                                   This sequence represents a modified Plasmodium falciparum gp190/MSP-1 (merzozite surface) protein. The gene encoding this protein has been stabilised by reducing the AT content of the nucleotide sequence. Such a protein is useful in vaccines against malaria or for producing monoclonal antibodies (for passive immunisation). The complete gp190 protein can now be produced outside the parasite and has, at least over extended regions, the native pattern of folding. Larger amounts of the protein can be produced recombinantly than would be possible using the parasites as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |:| : |:|||| : |:|||| : |:||| |:|| : |:||| |:||| : |:||| : |:|||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : ||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:|| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:||| : |:|||
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     246 TKKDMLGKLLSTGLVQIFPN---TIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHL 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  303 DEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPD 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         66 ESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       186 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 19; Length 1639;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             protein engineering; protein expression; codon usage;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MSP-1; merozoite surface protein; malaria; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; Pred. No. 3e-79; 54; Mismatches 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  60.8%; Score 1248.5; 61.0%; Pred. No. 3e-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1608 syplfdgifcssnflgiffllimlilysfi 1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  363 SYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Merozoite surface protein MSP-1-42.
                    Example 1; Fig 3c; 48pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY09372 standard; Protein; 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0085649
97US-0062592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98WO-US22226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                  Ë
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity
les 239; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  transgenic animal.
                                                                                                                                                                                                                                                                                                                                                               1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            31-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO9920774-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20-OCT-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-MAY-1998;
20-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29-APR-1999
                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY09372;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Loca
Matches
```

셤

ò g 요

ò

9 õ QQ Q

ò

셤

à

```
ä
                                                                                                                                                                                          NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD 245
                                                                                                                                                                                                                                                                                                                                             362
                                                                                                                                                                      ESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                              126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                           6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                               246 TKKDMLGKLLSTGLVQIFPN---TIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHL
                                                                                                                                                                                                                                                                                                                                                                                                                       DEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPD
                                                       25;
                   DB 20; Length 355;
56.4%; Score 1156.5; DB 20; Leuyu.
59.0%; Pred. No. 1.2e-73;
"...matches 73; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ą.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Merozoite surface protein MSP-1-42.
```

```
g
                                                   ò
                                                                                     셤
                                                                                                                      οy
                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The present sequence represents a 42 kba C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against target for the development of a vaccine against the plant along the sequence encoding MSP-1-42 has been modified (see AAX25586) compared to the native sequence (see AAX25587) such that 306 nucleotide positions have been replaced to lower the AT content (from 76 to 49.7%) and to eliminate 10 mRNA instability motifs while mainfaining the same protein amino acid sequence. These alterations allow MSP-1-42 to be expressed in mammalian cell culture and in transgenic mice. The invention to provides modified recombinant nucleic acid sequences and methods for increasing the mRNA levels and protein expression of proteins that are difficult to express in cell culture systems, mammalian cell culture systems or in transgenic aminals. The preferred difficult corporations such as parasites, bacteria and viruses that have DNA condant sequences of high AT content or which have mRNA instability motifs or rare codons relative to the recombinant expression system to be used. The invention allows expression of MSP-1 protein in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        66 ESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKY 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99
                                                                                                                                                                                                                                                                                                                                                                                         New modified recombinant nucleic acid sequences useful for producing malarial DNA vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             to be used. The invention allows expression of MSP-1 protein in the milk of transgenic animals, and also provides a DNA vaccine comprising a vector containing the altered MSP-1-42 sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 20; Length 355;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     73; Indels
                engineering; protein expression; codon usage;
 protein; malaria; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  56.4%; Score 1156.5; DB 59.0%; Pred. No. 1.2e-73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   54; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Disclosure; Fig 1; 43pp; English.
                                                                                                                                                                                                                                                                   (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                            98WO-US22225.
                                                                                                                                                                                                           98US-0085649.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 59.09
Matches 219; Conservative
MSP-1; merozoite surface
                                                                  Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                      WPI; 1999-302742/25.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                      Chen LH, Meade H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                355 AA;
                               transgenic animal
                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAX25586
                                                                                                     WO9920766-A2
                                                                                                                                                                          20-OCT-1998;
                                                                                                                                                                                                             5-MAY-1998;
                                                                                                                                                                                                                                20-OCT-1997;
                                                                                                                                       29-APR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           167
```

ò g δ g ŏ a

a ð

```
.;
                                                                                                     303 DEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPD 362
                                                                                                                                 6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This present sequence comprises a 42 kDa C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against plasmodium falciparum. The C-terminal end of the sequence is modified to include a 6xHis tag. A nucleic acid (see AAX56008) encoding MSP-1-42 has been modified according to a method of the invention in order to improve expression in mammalian cells and in transgenic animals by reducing the AT content and removing mRNA instability motifs. The invention allows expression of MSP-1-42 in the milk of transgenic animals, and also provides a DNA vaccine comprising a vector containing the altered MSP-1-42 nucleic
246 TKKDMLGKLLSTGLVQIFPN---TIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHL
                           56.4%; Score 1156.5; DB 20; Length 361; 59.0%; Pred. No. 1.2e-73; Live 54; Mismatches 73; Indels 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Modified malarial protein for use in anti-malarial vaccines
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MSP-1; merozoite surface protein; malaria; vaccine; protein engineering; protein expression; codon usage;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
.2e-73;
73;
                                                                                                                                                                                                                                                                                                                                                                                                AAY09373 standard; Protein; 361 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Merozoite surface protein MSP-1-42.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Example; Fig 2; 35pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US22226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98US-0085649.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     31-AUG-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity 59.0
tes 219; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Plasmodium falciparum.
                                                                                                                                                                                                             363 SYPLFDGIFCS 373
                                                                                                                                                                                                                                            WPI; 1999-288313/24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  transgenic animal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Meade H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          361 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               P-PSDB; AAX56009
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          W09920774-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20-0CT-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15-MAY-1998;
20-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              29-APR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY09373;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Chen LH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                    AAY09373
```

ë,

```
WO9920774-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
 88888888888888
                                                                                                                                                                         셤
                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                     186 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD 245
                                                                                                                 246 TKKDMLGKLLSTGLVQIFPN---TIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHL 302
ESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                     126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                    284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This present sequence comprises a 42 kDa C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against Plasmodium falciparum. The C-terminal end of the sequence is modified to include a KHIS tag. Nucleic acids (see AAX25586 and AAX25593) encoding MSP-1-42 have been modified according to a method
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           New modified recombinant nucleic acid sequences useful for producing malarial DNA vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 of the invention in order to improve expression in mammalian cells and in transgenic animals. The invention provides modified recombinant nucleic acid sequences and methods for increasing the mRNA levels and protein expression of proteins that are difficult to express in cell culture systems, especially mammalian cell
                                                                                      MSP-1; merozoite surface protein; malaria; vaccine; protein engineering; protein expression; codon usage;
                                                                                                                                                                                                                                                               Ā
                                                                                                                                                                                                                                                                                                                         Merozoite surface protein MSP-1-42.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Disclosure; Fig 2; 43pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GENZ ) GENZYME TRANSGENICS CORP
                                                                                                                                                                                                                                                              AAY05833 standard; Protein; 361
                                                                                                                                                                                                                                                                                                                                                                                                                                         98WO-US22225.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            98US-0085649.
97US-0062592.
                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                Plasmodium falciparum
                                                                                                                                                                                             363 SYPLFDGIFCS 373
                                                                                                                                                                                                       WPI; 1999-302742/25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Chen LH, Meade H;
                                                                                                                                                                                                                                                                                                                                                              transgenic animal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-PSDB; AAX25587
                                                                                                                                                                                                                                                                                                                                                                                                    W09920766-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                        20-OCT-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            15-MAY-1998;
20-OCT-1997;
                                                                                                                                                                                                                                                                                                     02-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                      29-APR-1999
                                                                                                                                                                                                                                                                                 AAY05833;
99
                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                              g
                g
                                     ò
                                                    g
                                                                          à
                                                                                           g
                                                                                                                 ò
                                                                                                                                                     ç
                                                                                                                                                                         g
                                                                                                                                                                                            ò
ò
```

```
NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             246 TKKDMLGKLLSTGLVQIFPN---TIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHL 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPD 362
                                                                                                                                                                                                                                                                                                                                                  25; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                65
culture systems or in transgenic animals. The preferred difficult protein candidates for expression are those derived from lower organisms such as parasites, bacteria and viruses that have DNA coding sequences of high AT content or which have mRNA instability motifs or rare codons relative to the recombinant expression system to be used. The invertion allows expression of MSP-1 in the milk of transgenic animals, and also provides a DNA vaccine comprising a vector containing the altered MSP-1-42 nucleic acid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVL
                                                                                                                                                                                                                                                                                                                                                                                                                            227 nhnnlltkflstgmv--fenlaktvlsnlldgnlggmlnisghgcvkkgcpgnsgcfrhl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /note= "Asn in native MSP-1-42 (N181Q mutation)"
278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /note= "Asn in native MSP-1-42 (N262Q mutation)"
                                                                                                                                                                                                                                                                                                     DB 20; Length 361;
                                                                                                                                                                                                                                                                                               Query Match . 56.4%; Score 1156.5; DB 20; Lengt Best Local Similarity 59.0%; Pred. No. 1.2e-73; Matches 219; Conservative 54; Mismatches 73; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MSP-1; merozoite surface protein; malaria; vaccine; protein engineering; protein expression; codon usage; transgenic animal; mutant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 'note= "beta-casein signal peptide"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Modified merozoite surface protein MSP-1-42.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY09374 standard; Protein; 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /note= "6xHis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SYPLFDGIFCS 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 197
                                                                                                                                                                                                                                361 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            31-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Synthetic.
                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY09374;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY09374
```

Chen LH,

```
The present sequence represents a 42 kDa C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against plasmodium falciparum. The sequence has been modified to include an N-terminal beta-casein signal peptide and an N-terminal GAHis tag. In addition, N181Q and N262Q mutations have been introduced to eliminate 2 N-glycosylation sites. These modifications allow the Calminate 2 N-glycosylation sites. These modifications allow the MSP-1-42 protein to be expressed in the milk of transgenic mice. The invention generally relates to modified recombinant nucleic acid sequences and methods for increasing the mRNA levels and call culture systems or in cell culture systems or in transgenic animals. The preferred difficult protein candidates for transgenic animals. The preferred difficult protein candidates for transgenic animals. The preferred difficult protein candidates for expression are those derived from lower organisms such as parasites, bacteria and viruses that have DNA coding sequences of high AT content or which have mRNA instability motifs or rare codons relative to the recombinant expression system to be used.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          126 LAKYKODLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 New modified recombinant nucleic acid sequences useful for producing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /note= "Asn in native MSP-1-42 (N181Q mutation)"
Misc-difference 278
                                                                                                                                                                                                                                                                                                          /note= "Asn in native MSP-1-42 (N262Q mutation)"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        56.0%; Score 1149.5; DB 20; Lengu.
58.8%; Pred. No. 4e-73;
... ....matches 75; Indels
                                  protein engineering; protein expression; codon usage; transgenic animal; mutant.
                                                                                                                                                                                    /note= "beta-casein signal peptide"
                  malaria; vaccine,
                                                                                                                                                    Location/Qualifiers
              MSP-1; merozoite surface protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Example; Fig 11; 43pp; English.
                                                                                                                                                                                                               371..376
/note= "6xHis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98US-0085649.
                                                                                                                                                                                                                                                                                                                                                                                                                              98WO-US22225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                         Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     malarial DNA vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1999-302742/25.
                                                                                                                                                                                                                                                   Misc-difference 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Meade H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 376 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-PSDB; AAX25593
                                                                                                                                                                                                                                                                                                                                                                                                                              20-OCT-1998;
                                                                                                                                                                                                                                                                                                                                                WO9920766-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-MAY-1998;
20-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                       29-APR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 218;
                                                                                                                 Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Chen LH,
                                                                                                                                                        Key
Peptide
                                                                                                                                                                                                               Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ||| ||:|||| || sekyksdldsikkyl------ndkqgenekylpflnnietlyktvn 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           66 ESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPD 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The present sequence represents a 42 kDa C-terminal portion of manalaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against Plasmodium falciparum. The sequence has been modified to include an N-terminal beta-casein signal peptide and an N-terminal faxHis and intending and N2620 mutations have been introduced eliminate 2 N-01ycosylation sites. These modifications allow the MSP-1-42 protein to be expressed in the milk of transgenic mice.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 20; Length 376;
                                                                                                                                                                                                                                                                                      Modified malarial protein for use in anti-malarial vaccines
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  56.0%; Score 1149.5; DB 20; Lengt|
58.8%; Pred. No. 4e-73;
ive 53; Mismatches 75; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Modified merozoite surface protein MSP-1-42.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY05834 standard; Protein; 376 AA
                                                                                                                                                  (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                          Example; Fig 11; 35pp; English.
                                                  98WO-US22226.
                                                                                       98US-0085649,
97US-0062592,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SYPLFDGIFCS 373
                                                                                                                                                                                                                             WPI; 1999-288313/24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
es 218; Conserv
                                                                                                                                                                                                                                                 P-PSDB; AAX56008
                                                                                                                                                                                        Meade
                                                  20-0CT-1998;
                                                                                                           20-OCT-1997;
                                                                                         15-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       02-AUG-1999
            29-APR-1999
```

Sequence

Query Match

Matches

qq

ò

186

ò

g

ò

82

3,

Length 376;

81

AAY05834;

EXTXCXD

363

ò

303

ö

g

QQ

ò

1284 idnilskieneyevlylkplagvyrslkkqlennvmtfnvnvkdilnsrfnkrenfknvl 1343

LAKYKODLESIKKVIKEEKEKFFSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV

ESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKV 125

99

Qγ

|:| : |:|||| ||||| : -----ndkqgenekylpflnnietlyktvn 1443 186 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD

1404 sekyksdldsikkyi-----

Db Qy Db

9

```
302
                                                                                                                                                                                                                                                                                                                                                                     DEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPD 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The sequence encoding the P195 protein of Plasmodium falciparum (AAN50530) and a peptide comprising at least one of its epitopes (see AAP50777) are claimed. Also claimed is a vaccine for inducing immunity to malaria comprising the novel peptide or P195 or a peptide comprising at least one epitope when derived from the new DNA sequence, together with a carrier.
                                                                                                                             246 TKKDMLGKLLSTGLVQIFPN---TIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHL
                                                                                                                                                                                                                                                    NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Cloned DNA sequence encoding plasmodium falciparum protein - useful for expressing the protein for use in vaccines against
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Riveros-Moreno V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence of the P195 protein of Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Malaria vaccine; epitope; antigen; immunogen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Lockyer M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAP50777 standard; Protein; 1654 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Odink K,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (WELL ) WELLCOME FOUNDATION LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 6; Fig 1; 51pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       84GB-0024340.
84GB-0004692.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              85EP-0301173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  85GB-0004429
sekyksdldsikkyi-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sandhu J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  syplfdgifcs 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                             SYPLFDGIFCS 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1985-224845/37.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1654 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAN50530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26-SEP-1984;
22-FEB-1984;
21-FEB-1985;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-FEB-1985;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30-SEP-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11-SEP-1985
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EP154454-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Holder A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAP50777;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        malaria
                                  142
                                                                                          186
                                                                                                                                                                                                                                                                                                                                         303
                                                                                                                                                                                                                                                                                                                                                                                                 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                             363
                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              H D D X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D 
                                                                                          ò
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                     ò
```

```
fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The non-natural variants have reduced affinity for at least 1 antibody capable of blocking a second antibody that inhibits the proteolytic cleavage of Plasmodium MSP-1_2, and has the same affinity for at least one third antibody that inhibits the proteolytic cleavage of plasmodium MSP-1_1_9. The non-natural variants of the present invention are useful for immunising a mammal against malaria, and can be used to treat malaria. The present sequence is MSP-133 protein.
                 302 LDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKP 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The present invention relates to non-natural variants of a C-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Uthaipibull C;
246 TKKDMLGKLLSTGLVQIFPN---TIISKLIEGKF-QDMLNISQHQCVKKQCPENSGCFRH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Novel variants of the C-terminal fragment of Plasmodium merozoite surface protein-1, useful as vaccines for treating or preventing malaria -
                                                                                                                                                                                                                                                                                                                                               Merozoite surface protein; protazoacide; vaccine; malaria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Syed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Morgan W,
                                                                                                                                       362 DSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                            AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Feeney J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Example 5; Fig 15; 126pp; English.
                                                                                                                                                                                                                           AAB37610 standard; Protein; 262
                                                                                                                                                                                                                                                                                                                  Merozoite surface protein-133.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99GB-0009072.
99US-0311817.
99CA-2271451.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20-APR-2000; 2000WO-GB01558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (MEDI-) MEDICAL RES COUNCIL
                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Holder A, Birdsall B,
                                                                                                                                                                                                                                                                                                                                                                           Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2001-015762/02.
N-PSDB; AAC68978.
                                                                                                                                                                                                                                                                                                                                                                                                       WO200063245-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            13-MAY-1999;
25-MAY-1999;
                                                                                                                                                                                                                                                                                    27-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                   26-OCT-2000
                                                                                                                                                                                                                                                        AAB37610;
                                                                                                                                                                                             RESULT
AAB37610
                                                                                                                                                                                                                                                        g
                                                            δλ
                                                                                        셤
                                                                                                                                               qq
```

**.**;

Gaps

MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVL 65

53.2%; Score 1091; DB 6; Length 1654; 55.2%; Pred. No. 3.6e-68; ive 60; Mismatches 90; Indels 26

Local Similarity 55.2 nes 217; Conservative

Matches 6

ò

Query Match

```
Disclosure; Fig 1B; 85pp; French
                                                                                 116 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P-PSDB; AAW22592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Barnwell JW,
                                                                                                                                                                                                                                                                                                                                                                                  WO9730159-A2
                                                                                                                                                                                                                                                                                                                                                                                                               14-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                             14-FEB-1996;
                                                                                                                                                                                                                                                    25-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                 21-AUG-1997
                                                                                                                                                                                                                                                                                                                Synthetic.
                                                                                                                                                                                                                                    AAW22592;
                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Roth C;
                                                                                                                                                                                                                                                                                                                                      Region
                                                                                                                                                                                                                                                                                                                                                             Region
                                                                                                                                                                                                        Π
                                                                                                                                                                                                                                                                                                                              Key
                                                                                                                                                                                                              AAW22592
 X \times C C C C C C C C X X C
                                                                                                                                                                                                                               q
                                                                                                                                                                                a
                                                                                                                                     ò
                                                                                                                                                                   õ
                                             'n.
                                                                                                                                                                                                                                                                                                                                                                                  sedneuce,
                                                                                                                                                                                                                                                                                                                                                         /note= "amino acids derived from P. falciparum MSP1 p19
                                                                                                                             LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                           ESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                             Gaps
                                                           MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Recombinant protein containing the merozoite surface protein-1 p19 fragment - useful in anti-malarial vaccines, diagnosis and protein purification
                                                                  NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD
                                            25;
                                                                                                                                                                                                                                                                                                                                                                                /note= "glycosylphosphatidylinositol anchoring
                              Length 262;
                                                                                                                                                                                                                                                                                               Plasmodium vivax; merozoite surface protein; MSP1; p19;
Plasmodium falciparum; malaria; vaccine; immunity; epitope.
                                            71; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ē
                              DB 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nato
                                                                                                                                                                                        227 nhnnlltkflstgmv--fenlaktvlsnlldgnlggml 262
                                                                                                                                                                                246 TKKDMLGKLLSTGLVQIFPN --- TIISKLIEGKFQDML 280
                                     .3e-37;
                                           51; Mismatches
                             Score 641.5;
Pred. No. 1.3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mendis K,
                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
                                                                                                                                                                                                                                      Ŗ
                                                                                                                                                                                                                                                                                                                                                                   fragment'
                                                                                                                                                                                                                                    AAW36103 standard; Protein; 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Longacre-Andre S,
                                                                                                                                                                                                                                                                                 PfMSP1(p19)A protein sequence.
                             31.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           PASTEUR.
NEW YORK STATE.
                                                                                                                                                                                                                                                                                                                                                                                                                             97WO-FR00290
                                                                                                                                                                                                                                                                                                                                                                                                                                            96FR-0001822
                                                                                                                                                                                                                                                                  (first entry)
                                    Best Local Similarity 47.1
Matches 131; Conservative
                                                                                                                                                                                                                                                                                                                     Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1997-425033/39.
N-PSDB; AAT94550.
       Ä
       262
                                                                                                                                                                                                                                                                                                                                                                                                                                                           (INSP ) INST
(UYNY ) UNIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Barnwell JW,
                                                                                                                                                                                                                                                                                                                                                                                               WO9730158-A2
                                                                                                                                                                                                                                                                                                                                                                                                                             14-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                            14-FEB-1996;
                                                                                                                                                                                                                                                                  25-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                              21-AUG-1997
                                                                                                                                                                                                                                                                                                                            Synthetic.
       Sequence
                                                                                                                                                                                                                                                  AAW36103;
                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Roth C;
                                                                                                                                                                                                                                                                                                                                                   Region
                                                                                                                                                                                                                                                                                                                                                                         Region
                                                                                       99
                                                                                                      67
                                                                                                                                                                                                                            AAW36103
X S
                                                                        a
                                                                                       à
                                                                                                                 ŏ
                                                                                                                                 셤
                                                                                                                                                             g
                                                                                                                                                                                                                                            쉽
                                                                                                                                                 ò
                                                                                                                                                                               ŏ
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          96..116
/note= "glycosylphosphatidylinositol anchoring sequence"
This is the amino acid sequence of a recombinant protein comprising amino acids 1613-1705 of the Plasmodium falciparum merozoite surface protein 1 (MSPI) 19 KD C-terminal fragment (p19), linked to a glycosylphosphatidylinositol membrane anchoring sequence. p19 is the C-terminal fragment of the 42 kD MSPI from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the p19 fragment provides a high level of protective immunity since it includes epitopes not presented in the p42 fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note= "amino acids derived from P. falciparum MSP1 p19
                                                                                                                                                                                                                                                                                                                                                                                                                                                         281 NISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADA 340
                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            nev
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  341 TCTEEDSGSSRKKITCECTKPDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Recombinant protein containing Plasmodium merozoite surface protein-1 p42 fragment - useful in antimalarial vaccines, also antibodies for diagnosis and protein purification
                                                                                                                                                                                                                                                                                                                                           Length 116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Plasmodium vivax; merozoite surface protein; MSP1; p19;
Plasmodium falciparum; malaria; vaccine; immunity; epitope.
                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ..
[14
                                                                                                                                                                                                                                                                                                                                        Score 615; DB 18;
Pred. No. 3.2e-36;
1; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mendis K,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Disclosure; Fig 1B; 85pp; French
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAW22592 standard; Protein; 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Longacre-Andre S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PfMSP1(p19)A protein sequence.
                                                                                                                                                                                                                                                                                                                                        30.0%;
ilarity 96.5%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              96FR-0001821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97WO-FR00291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (INSP ) INST PASTEUR.
(UYNY ) UNIV NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1997-425034/39.
                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 110; Conserv
```

Length 108;

œ

```
ö
This is the amino acid sequence of a recombinant protein comprising amino acids 1613-1705 of the Plasmodium falciparum merozoite surface protein 1 (MSP1) 19 kD C-terminal fragment (p19), linked to a glycosylphosphatidylinositol membrane anchoring sequence. p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the p19 fragment provides a high lavel of protective immunity since it includes epitopes not presented in the p42 fragment.
                                                                                                                                                                                                      281 NISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADA 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ပ်
                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Uthaipibull
                                                                                                                                                                                                                    Novel variants of the C-terminal fragment of Plasmodium merozoite surface protein-1, useful as vaccines for treating or preventing malaria -
                                                                                                                                                                                                                                                                    341 TCTEEDSGSSRKKITCECTKPDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                               ö
                                                                                                                                                    Length 116;
                                                                                                                                                                                                                                                                                                                                                                                                                                             Merozoite surface protein; protazoacide; vaccine; malaria.
                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Syed S,
                                                                                                                                                    Score 615; DB 18;
Pred. No. 3.2e-36;
1; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Morgan W,
                                                                                                                                                                               1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Feeney J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Example 5; Fig 15; 126pp; English.
                                                                                                                                                                                                                                                                                                                                             AAB37609 standard; Protein; 108
                                                                                                                                                                                                                                                                                                                                                                                                                      Merozoite surface protein-119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99GB-0009072.
99US-0311817.
99CA-2271451.
                                                                                                                                                    30.0%;
96.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20-APR-2000; 2000WO-GB01558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (MEDI-) MEDICAL RES COUNCIL
                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                             Matches 110; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Birdsall
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2001-015762/02
                                                                                                                                                                  Local Similarity
                                                                                                                  116 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAC68977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200063245-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      20-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                             27-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     26-OCT-2000
                                                                                                                    Sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Holder A,
                                                                                                                                                                                                                                                                                                                                                                    AAB37609;
                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                 AAB37609
                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                               88888888888888
                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                              8
```

The present invention relates to non-natural variants of a C-terminal fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The non-natural variants have reduced affinity for at least 1 antibody capable of blocking a second antibody that inhibits the proteolytic cleavage of Plasmodium MSP-1\_4\_2, and has the same affinity for at least one third antibody that inhibits the proteolytic cleavage of Plasmodium MSP-1\_4\_2. The non-natural variants of the present invention are useful for immunising a mammal against malaria, and can be used to treat malaria. The present sequence is MSP-119 protein.

108 AA

Sequence

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The present invention relates to non-natural variants of a C-terminal fragment of a Plasmodlum merozoite surface protein-1 (MSP-1). The non-natural variants have reduced affinity for at least 1 antibody capable of blocking a second antibody that inhibits the proteolytic cleavage of Plasmodlum MSP-14-2, and has the same affinity for at least one third antibody that inhibits the proteolytic cleavage of Plasmodlum MSP-12-9. The present sequence is the wild-type MSP-1 protein. This sequence was used to generate the variants of the present invention. The non-natural variants of the present invention. The mon-natural variants of the present invention are useful for immunising a mammal against malaria, and can be
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        281 NISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADA 340
                                                          272 IEGKFQDMLNISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCEENPNPTCNE 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ΰ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Uthaipibull
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Plasmodium merozoite
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     of Plasmodium merozoito treating or preventing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 96;
                                                                                                                                                                                                                                                                                                                                           Merozoite surface protein; protazoacide; vaccine; malaria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                     64 nnggcdadakcteedsgsngkkitcectkpdsyplfdgifcsssn 108
                                                                                                                    332 NNGGCDADATCTEEDSGSSRKKITCECTKPDSYPLFDGIFCSSSN 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Syed
Score 529; DB 22;
Pred. No. 3.3e-30;
4; Mismatches 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 528; DB 22;
Pred. No. 3.3e-30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Morgan W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Novel variants of the C-terminal fragment surface protein-1, useful as vaccines for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Example 2; Page 48; 126pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Holder A, Birdsall B, Feeney J,
                                                                                                                                                                                                                        AAB37608 standard; protein; 96 AA
 25.8%;
88.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99GB-0009072.
99US-0311817.
99CA-2271451.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Match 25.7%;
Local Similarity 94.8%;
nes 91; Conservative
                                                                                                                                                                                                                                                                                                              Merozoite surface protein-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (MEDI-) MEDICAL RES COUNCIL
                                                                                                                                                                                                                                                                                                                                                                                                                                                             20-APR-2000; 2000WO-GB01558
                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             treat malaria.
                                                                                                                                                                                                                                                                                                                                                                       Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2001-015762/02.
            Local Similarity
nes 93; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ¥,
                                                                                                                                                                                                                                                                                                                                                                                                    WO200063245-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-MAY-1999;
25-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          20-APR-1999;
                                                                                                                                                                                                                                                                                  27-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                                                               26-OCT-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Si
Matches 91;
                                                                                                                                                                                                                                                     AAB37608;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             used to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      malaria
                  Best Loca
Matches
                                                                                                                                                                                                          AAB37608
                                                                                                                                                                                                                                         ò
                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
```

σ

ö

Indels

5,

Mismatches

1;

Conservative

```
90;
                                                                                                                                                                                                                                                                                                Chimeric
Chimeric
                                                                                                                                                                                   AAW36102;
                                                                                                                                                                                                                                                                                                                                                 Peptide
                                                                                                                                                                                                                                                                                                                                                                           Protein
                                                                                                                                                                                                                                                                                                                                                                                                                          Region
                                                                                                                                                                                                                                                                                                                                                                                                    Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Region
                                                                                                                                     RESULT 15
Matches
                                                                                                                                                                                                                                                                                                                                        Key
                                                                                                                                               AAW36102
                                               g
                                                                                              Db
                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This is the amino acid sequence of a chimeric protein comprising amino acids 1-32 of the Plasmodium vivax merozoite surface protein 1 (MSP1) linked to amino acids Asn1613-Ser1705 which correspond to the 19 kD c-terminal fragment (P19) of MSP-1 from Plasmodium falciparum. p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the p19 fragment provides a high level of protective immunity since it includes epitopes not presented in the p42 fragment.
                                                                                                                                                                                                                                                                                                                                                                           Ç
                                                                                                                                                                                                                                                                                                                                                                       /note= "encoded by restriction enzyme sequence used
  create the chimeric sequence"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Recombinant protein containing Plasmodium merozoite surface protein-1 p42 fragment - useful in antimalarial vaccines, also new antibodies for diagnosis and protein purification
                                                                                                                                                                                                                                                                                                                                                                                                             /note= "derived from P. falciparum C-terminal p19
fragment of MSP1"
                                                                                                                                                                                              Chimeric, Plasmodium vivax; merozoite surface protein; MSP1; p19; Plasmodium falciparum; malaria; vaccine; immunity; epitope.
                                                                                                                                                                                                                                                                                                                                               /note= "derived from P. vivax MSP1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Nato F;
          341 TCTEEDSGSSRKKITCECTKPDSYPLFDGIFCSSSN 376
                       Mendis K,
                                                                                                                                                                                                                                                                                            /note= "signal peptide"
                                                                                                                                                                                                                                                                                                                       /note= "mature protein"
                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                             AAW22593 standard; Protein; 127 AA.

    Plasmodium vivax.
    Plasmodium falciparum.

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Disclosure; Fig 1C; 85pp; French
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sarnwell JW, Longacre-Andre S,
                                                                                                                                                                    PfMSP1(p19)S protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PASTEUR.
NEW YORK STATE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 97WO-FR00291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        96FR-0001821
                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                          ..127
                                                                                                                                                                                                                                                                                                                                                                                                 35..127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1997-425034/39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      127 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      P-PSDB; AAW22592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (INSP ) INST
(UYNY ) UNIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                WO9730159-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 14-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        14-FEB-1996;
                                                                                                                                             25-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21-AUG-1997.
                                                                                                                     AAW22593;
                                                                                                                                                                                                                                  Chimeric
                                                                                                                                                                                                                                              Chimeric
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                  Peptide
                                                                                                                                                                                                                                                                                                            Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Roth C;
                                                                                                                                                                                                                                                                                                                                   Region
                                                                                                                                                                                                                                                                                                                                                             Region
                                                                                                                                                                                                                                                                                                                                                                                                 Region
                                                                                                                                                                                                                                                                         Key
          õ
                               g
```

Length 127;

Score 521; DB 18; Pred. No. 1.5e-29;

25.4%; 93.8%;

Query Match Best Local Similarity

```
This is the amino acid sequence of a chimeric protein comprising amino acids 1-32 of the plasmodium vivax merozoite surface protein 1 (MSP1) linked to amino acids Asni613-Ser1705 which correspond to the 19 kD c-terminal fragment (pl9) of MSP-1 from Plasmodium falciparum. pl9 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the pl9 fragment provides a high level of protective immunity since it includes epitopes not presented in the p42 fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ţ
/note= "encoded by restriction enzyme sequence used
  create the chimeric sequence"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Recombinant protein containing the merozoite surface protein-1 p19 fragment - useful in anti-malarial vaccines, diagnosis and protein purification
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note= "derived from P. falciparum C-terminal p19
fragment of MSP1"
                                                                                                                                                                                                                                                                  Chimeric; Plasmodium vivax; merozoite surface protein; MSP1; p19;
                                                                                                                                                                                                                                                                              Plasmodium falciparum; malaria; vaccine; immunity; epitope.
                                                                                                                                                                                                                                                                                                                                                                                                                                      /note= "derived from P. vivax MSP1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nato F;
                                                       338 ADATCTEEDSGSSRKKITCECTKPDSYPLFDGIFCS 373
                                                                    92 adakcteedsgsngkkitcectkpdsyplfdgifcs 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mendis K,
                                                                                                                                                                                                                                                                                                                                                                            /note= "signal peptide"
                                                                                                                                                                                                                                                                                                                                                                                                           /note= "mature protein"
                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                       A

    Plasmodium vivax.
    Plasmodium falciparum.

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Disclosure; Fig 1C; 85pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Barnwell JW, Longacre-Andre S,
Roth C;
                                                                                                                                                  AAW36102 standard; Protein; 127
                                                                                                                                                                                                                                       PfMSP1(p19)S protein sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   96FR-0001822.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         97WO-FR00290.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (INSP ) INST PASTEUR.
(UYNY ) UNIV NEW YORK STATE.
                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                           ..127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1997-425033/39.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-PSDB; AAT94549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO9730158-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        14-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14-FEB-1996;
                                                                                                                                                                                                           25-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21-AUG-1997
```

```
SQ Sequence 127 AA;
```

```
;
0
                                                             Gaps
Query Match 25.4%; Score 521; DB 18; Length 127; Best Local Similarity 93.8%; Pred. No. 1.5e-29; Matches 90; Conservative 1; Mismatches 5; Indels (
```

q ٥y

Search completed: August 8, 2001, 12:31:55 Job time: 163 sec

```
NAME/KEY: Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JS-08-290-919-3
                                                114.5
1111.5
1111.5
1111.5
1111.1
111.1
111.1
111.1
110.5
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
108.1
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 1, Applisequence 12, Applisequence 12, Applisequence 11, Applisequence 11, Applisequence 115, Appli
                                                                                                                                                                                                                                                                                                 (without alignments)
235.763 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 120, App
Sequence 120, App
Sequence 1, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 AISVTMDNILSGFENEYDVI......SNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 2,
Sequence 6,
Sequence 6,
                                                                                                                                                                                                                                                         ; Search time 34.41 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-290-919-4
US-08-290-919-12
US-08-290-919-12
US-08-290-919-11
US-08-75-587-186
US-08-257-073-5
US-08-480-640A-193
US-08-257-073-5
US-08-295-802-115
US-08-666-968C-58
US-08-666-968C-193
US-08-666-968C-193
US-08-688-237A-115
US-08-488-237A-193
US-08-488-237A-193
US-08-973-03077-3
US-08-188-237A-193
US-08-188-237-193-193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          197339 seqs, 20590346 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                                    August 8, 2001, 12:29:12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-500-376-2
2062
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        282
277
277
277
271
271
286.5
1123
1123
113
119
119
119
110
111
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
11
                                                                                                                                                                                                                                                                                                                                                                                                                             Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Scoring table:
                                                                                                                                                                        OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Database
                                                                                                                                                                                                                                                         Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Result
```

Š

```
sequence 1, Appl
Sequence 14, Appl
Sequence 14, Appl
Sequence 14, Appl
Sequence 14, Appl
Sequence 4, Appl
Sequence 44, Appl
Sequence 52, Appl
Sequence 2, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: HOLDER, ANTHONY A.
APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: UACCINE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE S.S.
ADDRESSE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COMPUTER READABLE FORM:
MEDITHY TYPE: Floppy disk
COMPUTER READABLE FORM:
MEDITHY TYPE: PLOPPY disk
COMPUTER: DEPOSY disk
COMPUTER: DEPOSY disk
COMPUTER: DEPOSY disk
COMPUTER: PREADABLE FORM:
PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPOTER: I LEM PC COMPOTER: COMPOTER: COMPOTER: Patentin Release #1.0, Version #1.25

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919

FILING DATE: U4-OCT-1994

CLASSIFICATION ADATA:
APPLICATION NUMBER: GB 9203821.5

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00367

FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: RCKULLS, PAUL N.
REGISTRATION NUMBER: 16,773

RELEFRANCE/POCKET NUMBER: 16,773

RELECOMMUNICATION INFORMATION:
TELEFRANCE/POCKET NUMBER: 16,773

RELECOMMUNICATION INFORMATION:
TELEFRANCE/POCKET NUMBER: 212242/HCM/MJL/6BC8/
TELEFRANCE/POCKET NUMBER: 212242/HCM/MJL/6BC8/
TELEFRANCE/POCKET NUMBER: 22.9944

TELEFRANCE/POCKET NUMBER: 23.5000

TELEFRANCE/POCKET NUMBER: 23.5000
                                               US-08-540-804-14
US-08-510-804-14
US-08-510-399-14
US-09-104-324B-4
US-09-116-324B-4
US-09-755-587-184
US-09-755-587-184
US-09-755-587-16
US-09-755-587-16
US-09-044-908-2
US-09-044-908-2
US-08-335-700-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 3, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
```

ö

Gaps

ö

Indels

```
325 PNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSN 376
                                                                   2 PNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPDSYPLFDGIFCSSN 53
                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P. STREET: 1100 NEW YORK AVENUE, N.W. CITY: WASHINGTON STATE: D.C.
  1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                          Sequence 12, Application US/08290919
Patent No. 5720959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 2, Application US/08290919; Patent No. 5720959; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE/DOCKET NUMBER: 21
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEPAN: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: HOLDER, ANTHONY A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 13.4%;
Best Local Similarity 97.9%;
Matches 47; Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 48 amino acids
    49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; MOLECULE TYPE: peptide US-08-290-919-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: amino acid
STRANDEDNESS: Si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY:
                                                                                                                                                  RESULT 3
US-08-290-919-12
    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    op
Op
                                             ò
                                                                                                                                                                        ö
                                                                                                                                                                                                                325 PNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSN 376
                                                                                                                                                                                                                                     2 PNPTCNENNGGCDADARCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSN 53
                                                                                                                                                                                                                                                                                                                                                     Sequence 4. Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.,
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESSS:
ADDRESSEE: CUSHMAN DARRY & CUSHMAN, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
                                                                                                                               Length 53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 53;
                                                                                                                                                                      0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIble
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
                                                                                                                            Score 301; DB 1; Pred. No. 4.9e-16;
                     /label= X
/note= "X = M and N, or N"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    - M and N, or N"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 282; DB 1;
Pred. No. 1.2e-14;
                                                                                                           14.6%; Scor.
100.0%; Pred. No. ...
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERNEC/CDCKET NUMBER: 212242/HCM/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: US/08/290,919
FILING DATE: 04-OCT-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /label= X
/note= "X
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 4:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME/KEY: Modified-site
                                                                                                                    Query Match
Best Local Similarity 100.(
Matches 52; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            53 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: protein
; LOCATION: 1
; OTHER INFORMATION:
; OTHER INFORMATION:
US-08-290-919-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WASHINGTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                RESULT 2
US-08-290-919-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-290-919-4
                                                                                                                                                                                                                ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
GENERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      281 NISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCVENPNPT 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 NISQHQCVKKQCPQNSGCFRHLDEREECKCLLNYKQEGDKCVENPNPT 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE PATENTIAN BATA:

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/290,919

FILING DATE: US/08/290,919

FILING APPLICATION DATA:

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATE: 22-FEB-1992

PRIOR APPLICATION DATE: C22-FEB-1993

APPLICATION DATE: C22-FEB-1993

ATORNEY/AGENT INFORMATION:

NAME: REGISTRATION NUMBER: POT/GB93/00367

FILING DATE: Z2-FEB-1993

ATORNEY/AGENT INFORMATION:

NAME: REGISTRATION NUMBER: 16,7733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             212242/HCM/MJL/6BC8/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 277; DB 1;
Pred. No. 2.6e-14;
1; Mismatches 0
```

```
APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13.1%; Score 271; DB 1; Length 48; ilarity 97.9%; Pred. No. 7.2e-14; Conservative 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          282 ISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCVENPNPT 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 ISQHQCVKKQCPQNSGCFRHLDEREECKCLLNYKQEGDKCVENPNPT 48
                                      212242/HCM/MJL/6BC8/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CTHER INFORMATION: /label= x orHER INFORMATION: /note= "X = M and N, or US-08-290-919-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P. STREET: 1100 NEW YORK AVENUE, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 11, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 212
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SED ID NO: 1:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME/KEY: Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGIH: 48 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STREET: 1100 NEW CITY: WASHINGTON STATE: D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
hes 46; Conserv
         WASHINGTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-290-919-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
BLACKMAN, MICHAEL J.
CHAPPEL, JONATHAN A.
VENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS: 3
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13.3%; Score 274; DB 1; Length 48; 100.0%; Pred. No. 4.3e-14; Live 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 282 ISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCVENPNPT 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 ISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCVENPNPT 48
                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION 1435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION DATA:
APPLICATION DATA:
APPLICATION UNBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
FELECOMMUNICATION INFORMATION:
TELECOMMUNICATION OF 822-0944
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
                                                                                                                                                                                                                                                                                                  COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARRE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OTHER INFORMATION: /label= X
OTHER INFORMATION: /note= "X - M and N, or
                                                                                                     CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1100 NEW YORK AVENUE, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 1, Application US/08290919
Patent No. 5720959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HOLDER, ANTHONY A
                                                                 VACCINE
                                                                                                                                                                                                                                                                                     Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME/KEY: Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 13.3°
Best Local Similarity 100.
Matches 47; Conservative
                                        TITLE OF INVENTION: IMP
TITLE OF INVENTION: VACC
NUMBER OF SEQUENCES: 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                            COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: linear
                                                                                                                                                                        CITY: WASHINGTON
                                                                                                                                                                                             STATE: D.C. COUNTRY: U.S.A. ZIP: 20005-3918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GENERAL INFORMATION:
APPLICANT: HOLDER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-290-919-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 5
US-08-290-919-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

ö

ö

```
FILING DATE: 14-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9617961.9
FILING DATE: 28-AUG-1996
ATTORNEY,AGENT INFORMATION:
NAME: Renneth D 51bley
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5405-135
REPERENCE CHARACTERISTICS:
LENGTH: 1494 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STATE: California
COUNTRY: United States of America
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 30, Application US/07728215 Patent No. 5962643 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92122
                                                                                                                                                                                                                                                                                                                                US-08-755-587-186
                                                                                                                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Futreal, Phillip A
APPLICANT: Wooster, Richard F
APPLICANT: Moster, Alan
APPLICANT: Shworth, Alan
APPLICANT: Stratton, Michael R
TITLE OF INVENTION: Materials and methods relating to the
TITLE OF INVENTION: identification and sequencing of the BRCA2 cancer
TITLE OF INVENTION: susceptibility gene and uses thereof.
CORRESPONDENCE: 222
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   278 DMLNIS-QHQCV-KKQCPENSGCFRHLDEREECKCLLNYKQ-EGDKCVENPNPTCNENNG 334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3: Bell Seltzer Park & Gibson
310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 106;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/755,587
FILING DATE: 25-NOV-1996
                                                                                 PRIOR APPLICATION 435
PRIOR APPLICATION 435
PRIOR APPLICATION 1948
APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATTORNEY AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 12,773
REFERENCE/DOCKET NUMBER: 21242/HCM/MJL/6BC8/
TELEPHONE: (202) 861-3000
TELEFONE, 1202 861-3000
TELEFONE, 6714627 CUSH
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 106 amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 335 GCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSN 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               65 GCDPTASCQNAESTENSKKIICTCKEPTPNAYYEGVFCSSSS 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1,2e-13;
has 32;
PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12.9%; Score 266.5;
49.0%; Pred. No. 4.26
tive 17; Mismatches
                      CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919
FILLING DATE: 04-OCT-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: GB 9523959.6 FILLING DATE: 23-NOV-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 186, Application US/08755587 Patent No. 6045997 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 50; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STREET: 310 UC
CITY: Raleigh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-755-587-186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 us-08-290-919-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

```
985 YKQTENLKISDHA-----SQKVDVHEN-TEKETAK------KPTMYTNQSTYSAIENSP 1031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S12 WTEEELNNFSDSLNSELLPGIDIKKTDISNHEVIE-----NTERKDKITKESLIGTENI 865
                                                                                                                                                                                                                                                                                                                                                                                                                                                            109 -----ENDIKFAQEG--ISYYE---KVLAKYKDDLESIKKVIKEEKEKFPSSPPTTP 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       205 EKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDDTKKDMLGKLLSTGLVQNFP 264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         265 NTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCVEN 324
                                                                                                                                                                                                                                760 VTMSNKQQLTANTGQNIKDFDTFYL----SFQTASRKNIRVSRESLNKARSLLNQK 811
                                                                                                                                                                                                                                                                                                         55 LKKRKY--FLDVLESDLMQFKHIS----SNEYIIEDSFKLLNSEQKNTLLKSYKYIKESV 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       156 PSPAKTDEQKKESK-----FLPFLTNIETLYNNL-VNKIDDYLINLKAKINDCNV 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             925 HRGAKMSKDREECKDGRELACGTTEITTPEYEETHSSLEKKKVSNEIAALRPRLLSDNL 984
                                                                                                                                                          4 VTMDNILSGFEN-----EYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLN----DILNSR 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: Shepard, Dean
APPLICANT: Shepard, Dean
APPLICANT: Quaranta, Vito
APPLICANT: Quaranta, Vito
APPLICANT: Quaranta, Vito
APPLICANT: Quaranta, Vito
APPLICANT: Quaranta, Nober.
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCE: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1085 -NPSC----GNSSNSAITENDKNHLSEKQGSTYLSNSTMSNSYSYHPG-FCHSS 1132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        325 PNPTCNENNGGCDADAKCTEED ---- SGSNGKKITCECTKPDSYPLFDGIFCSSS 375
                                                                                 Indels
Ouery Match 6.3%; Score 129.5; DB 3; Best Local Similarity 22.9%; Pred. No. 0.17; Matches 95; Conservative 68; Mismatches 167;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Patentin Palace
```

```
KVL-----AKYKDDLE------SIKKVIKEEKEK-FPSSPPTTPPSPAKTD 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KGLPEPTVTNEETVEELKKGILDMGIKLLFSKVKSLLKKLKNKIFPKKKEDNQAVDTKSM 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EEPK-VKAQPALRGVEPTEDSNIMNSINNVMDEIDFFEKELIENNNTPNVVPPTQSKKKN 527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------EKDEAHVKITKLS-DLKAIDDKIDLFKNHNDFE 236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          528 KNETVSGMDENFDNHPENYFKEEYYYDENDDMEVKVKKIGVTLK----KFEPLKNGNVSE 583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              584 TIKLIHLGNKDKKHIEAINNDIQIIKQELQAIYNE--LMNYTNG--NKNIQQIFQONILE 639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----CPENSGCFRHLDEREECKCLLNYKQEGDK 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  640 NDVLNQETEEEMEKQVEAITKQIEAEVDALAPKNK-----EEEEKEKEKEKEKEKEKEK 692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 --KLINDDTK --KDMLGKLLSTGLVQNFPNTIISKLIEGKFQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    47 LNDILNSRLKKR---KYFLDVLESDLMQFKHISSNEYI-----IEDSFKLLNSEQ-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EQKKESKFLPFL-----TNIETLYNNLVNKIDDYLINLKAKINDCNV---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 1; Length 740; 0.21;
                                                                                                                                                   MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTMARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    53; Mismatches 110;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6.0%; Score 123; 22.9%; Pred. No. 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: Frommer, William S.
REGISTATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2570
TELECOMMUNICATION INFORMATION:
  530 Fifth Avenue, 25th Floor
                                                                                                                                                                                                                                                                                                                           CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/075,783
FILING DATE: 11-UNN-1993
PRIOR APPLICATION NUMBER: US 07/852,305
APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE: 18-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/672,183
                                                                            UNITED STATES OF AMERICA
                                                                                                                                                                                                                                                                              APPLICATION NUMBER: US/08/257,073
FILING DATE: 09-JUN-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEEX: 425066 CURTMS
INFORMATION FOR SEO ID NO: 5:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE: 20-MAR-1991
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    96; Conservative
                                                                                                                              COMPUTER READABLE FORM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
Matches 96; Conserv
                            CIII.
STATE: New
COUNTRY: UNI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE:
FRAGMENT TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       94 -KNTL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-257-073-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                190 KTVMPYISTTPAKLRNPCTSEQNCTTPFSYKNVLSLTNKGEVFNELVGKQRISGNLDSPE 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      104 IKESVENDIKFAQEGISYYEKVLAKYKDDLESIKKV---IKEEKEKFPSS------ 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            138 -AEDYPIDLYYLMD-LSY-----SMKDDLENVKSLGTDLMNEMRRITSDFRIGFGSFVE 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              151 ---PPTTPPSPAK----TDEQKKESKF----LPFLTNIETLYNNLVNK-----ID--- 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----KAKI-----ND--CNVEKDE-- 208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               250 GGFDAIMQVAVCGSLIGWRNVTRLLVFSTDAGFHFAGDGKLGGIVLPNDGQCHLENNMYT 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 209 ------AHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDDTKKDMLGKLLSTG 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             310 MSHYYDYPSIAHL-VQKLSE----NNIQTIFAVTEEFQPVYKELKNLIPKSAVGTLSANS 364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 365 SNVIQLIIDAYNSLSSEVILENGKLSECVTISYKSYCKNGVNGTGENGRKCSNISIGDEV 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  425 QFEISITSNKCPKKDSDSFKIRPLGFTEEVEVILQYICECECQSEGIPESPKCHEG-NGT 483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         48 NDILNSR----LKKRKYFLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKY 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         78 DDIENPRGSKDIKKNKNVTNRSKGTAEKLKPEDIHQIQPQQLVLRLRSGEPQTFTLKFKR 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----CNENNGG----C-----DADAKCTEEDS----GSNGKKI--TCECTKPD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   290 -----KQCP-ENSGCFR-----HLDERE-----ECKCLLNYKQEGDKCVENPNPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .; Score 123.5; DB 2; Length 798;
.; Pred. No. 0.21;
67; Mismatches 146; Indels 179;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
APPLICANT: de Taisne, Charles
APPLICANT: Tine, John A.
TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
NUMBER OF SEQUENCES: 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----LVQNFPNTISKLI--EGKFQDMLNISQHQCVK--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Curtis, Morris & Safford, P.C
                                                                                                         ATTORNEY AND ALTONOMY

AMARE: Campbell, Cathryn A. REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P31 8717

TELECOMONICATION INFORMATION:
TELEPHONE: (619) 535-8949

INFORMATION FOR SEO ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 798 amino acids
TYPE: AMINO ACID
STRANDEDNESS: Single
                     APPLICATION NUMBER: US/07/728,215 FILING DATE: 19910711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 5, Application US/08257073 Patent No. 5766597
                                                                                                 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 20.8<sup>†</sup>
Matches 103; Conservative
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : :: | || || 544 INELYSGKFCECDNF 558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        363 SYPLFDGIFCSSSNF 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, M
                                                                                                                                                                                                                                                                                                                                                                                                    ; TOPOLOGY: linear
US-07-728-215-30
                                                                      CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             190 ---DYLINL--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 9
US-08-257-073-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              å
```

22;

Gaps

```
393 SATIK---SKLLALRAVKNWKCYSLTNVSMYKKIKGVIVMDMVDYISTNILKYHKQLYDK 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    185 DAFHIPPTHISLRSL-RDINRIIELLKKYPNNNIIDYISDSIKSNSSFIHILHMIISNMF 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               244 PAIIPSVNDFISTVVDKDRLINMYGIKCVAMFSYDINMIDLESLDDSDYIFIEKNISIYD 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     244 DDTKKDMLGKLLSTGLVQNF----PNTIISKLIEGKF-QDMLNISQHQCVK--KQCPEN 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      71 QFKHISSNEYI--IEDSFKLLNSEQ-KNTLLKSYKY----IKESVENDIKFAQEGISYYE 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      124 KVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNN 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLIN 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ENNKVEEVLIHID-----NVSKNNK----LSLSDISSLMDQFRL----NPCTIRNILLS 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          244 DDTKKDMLGKLLSTGLVQNF----PNTIISKLIEGKF-QDMLNISQHQCVK--KQCPEN 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18 DVIYLKPLAGVYRSLKKQIEKNIFTF----NLNLNDILNSRLKKRKYFLDVLE---SDLM 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 677;
                                                                                                          SGCFRHLDEREECKC-----LLNYKQEGDKCVENPNPTCNE 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA: APPLICATION NAMBER: US/08/4R0 6447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5.8%; Score 119; DB 3; L6
22.2%; Pred. No. 0.37;
:ive 65; Mismatches 137;
                                                                                                                                                                                                                                                                                                                                     Recombinant Swinepox Virus 225
                                                                                                                                                                                                                                                                                                                                                                                                         E: John P. White
1185 Avenue of the Americas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US/08/480,640A
07-JUN-1995
                                                                                                                                                                                                                               Sequence 193, Application US/08480640A
Patent No. 6033904
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David E.
TITLE OF INVENTION: Recombinant Swi
NUMBER OF SEQUENCES: 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3: Floppy disk
IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEFRX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 193:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELECOMMUNICATION INFORMATION TELEPHONE: (212) 278-0400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     677 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 22.2
Matches 76; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   protein
                                                                                                                                                                                                                                                                                                                                                                                   CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                 New York
: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDIUM TYPE:
                                                                                                                                                                                            RESULT 11
US-08-480-640A-193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-480-640A-193
                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                               STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                    296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            184
                                       g
                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               185 DAFHIPPTHISLRSL-RDINRIIELLKKYPNNNIIDYISDSIKSNSFIHILHMIISNMF 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 244 PAIIPSVNDFISTVVDKDRLINMYGIKCVAMFSYDINMIDLESLDDSDYIFIEKNISIYD 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    304 ---VKCRDFANMIRDKVKREKNRILTT------KCEDIIRYIKLFS-----KNRIND 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            184 LVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLIN 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         71 QFKHISSNEYI--1EDSFKLLNSEQ-KNTLLKSYKY----IKESVENDIKFAQEGISYYE 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            124 KVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNN 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18 DVIYLKPLAGVYRSLKKQIEKNIFTF----NLNLNDILNSRLKKRKYFLDVLE---SDLM 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 3; Length 677;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; Pred. No. 0.37;
65; Mismatches 137; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                               SOFTWARE: Patentin Release #1.0, Version #1.25
CUBRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,640A
FILING DATE: 07-JUN-1995
                                                                                                   GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David E.
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 119;
                                                                                                                                                                                                                                   ADDRESSEE: John P. White STREET: 1185 Avenue of the Americas
                                                           Sequence 115, Application US/08480640A Patent No. 6033904
                                                                                                                                                                                                                                                                                                                                                                                   E: Floppy disk
IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION: 424
ATTORNEY AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM: Swinepox virus
STRAIN: Kasza
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 677 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 22.28
Matches 76; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         double
                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INDIVIDUAL ISOLATE:
                                                                                                                                                                                                                                                                                                                                                                                                                           OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: DNA
                                                                                                                                                                                                                                                                                New York
New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IMMEDIATE SOURCE:
CLONE: 515-85.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POSITION IN GENOME MAP POSITION: -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ANTI-SENSE: NO
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                          USA
                                                                                                                                                                                                                                                                                                                                                                                        MEDIUM TYPE:
                      RESULT 10
US-08-480-640A-115
                                                                                                                                                                                                                                                                                                                                           10036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-480-640A-115
                                                                                                                                                                                                                                                                                                                     COUNTRY:
```

18;

Gaps

64;

g

ò

g ò

ŏ

à g

```
393 SATIK---SKLLALRAVKNWKCYSLTNVSMYKKIKGVIVMDWVDYISTNILKYHKQLYDK 449
     184 LVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLIN 243
                                                      347 ENNKVEEVLIHID-----NVSKNNK----LSLSDISSLMDQFRL----NPCTIRNILLS 392
                                                                                                              244 DDTKKDMLGKLLSTGLVQNF----PNTIISKLIEGKF-QDMLNISQHQCVK--KQCPEN 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                185 DAFHIPPTHISLRSL-RDINRIIELLKKYPNNNIIDYISDSIKSNSSFIHILHMIISNMF 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 71 QFKHISSNEYI--IEDSFKLLNSEQ-KNTLLKSYKY----IKESVENDIKFAQEGISYYE 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    244 PAIIPSVNDFISTVVDKDRLINMYGIKCVAMFSYDINMIDLESLDDSDYIFIEKNISIYD 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          124 KVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNN 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              304 ---VKCRDFANMIRDKVKREKNRILTT------KCEDIIRYIKLFS-----KNRIND 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       184 LVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLIN 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          347 ENNKVEEVLIHID-----NVSKNNK----LSLSDISSLMDQFRL----NPCTIRNILLS 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   244 DDTKKDMLGKLLSTGLVQNF----PNTIISKLIEGKF-QDMLNISQHQCVK--KQCPEN 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        393 SATIK---SKLLALRAVKNWKCYSLTNVSMYKKIKGVIVMDMVDYISTNILKYHKQLYDK 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18 DVIYLKPLAGVYRSLKKQIEKNIFTF----NENLNDILNSRLKKRKYFLDVLE---SDLM: 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 5.8%; Score 119; DB 4; Length 677; Best Local Similarity 22.2%; Pred. No. 0.37; Matches 76; Conservative 65; Mismatches 137; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              450 MSTFEYKRDIKSCKCSICSDSITHHIYETTSCI-NYKSTDND 490
                                                                                                                                                                                                                                                          450 MSTFEYKRDIKSCKCSICSDSITHHIYETTSCI-NYKSTDND 490
                                                                                                                                                                                                                     296 SGCFRHLDEREECKC-----LLNYKQEGDKCVENPNPTCNE 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         296 SGCFRHLDEREECKC-----LLNYKQEGDKCVENPNPTCNE 331
                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 58, Application US/08686968C
Patent No. 6221361
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
TITLE OF INVENTION: Recombinant Swinepox Virus
FILE REFERENCE: 39119-H/JML
CURRENT APPLICATION NUMBER: US/08/686,968C
CURRENT FILING DATE: 1996-07-25
NUMBER OF SEQ ID NOS: 231
SOFTWARE: Patentin Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 193, Application US/08686968C
Batent No. 6221361
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David E.
TITLE OF INVENTION: Recombinant Swinepox Virus
FILE REFERENCE: 39119-H/JML
CURRENT APPLICATION NUMBER: US/08/686,968C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ) ORGANISM: Swinepox virus
US-08-686-968C-58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-686-968C-193
                                                                                                                                                                                                                                                                                                                                                                                       US-08-686-968C-58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: PRT
                                                                                                            δλ
                                                                                                                                                               q
                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | :| |: :| || : :| || 304 ---VKCRDFANMIRDKVKREKNRILTT-------KCEDIIRYIKLFS-----KRRIND 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     71 QFKHISSNEYI--IEDSFKLLNSEQ-KNTLLKSYKY----IKESVENDIKFAQEGISYYE 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      244 PAIIPSVNDFISTVVDKDRLINMYGIKCVAMFSYDINMIDLESLDDSDYIFIEKNISIYD 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             124 KVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNN 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18 DVIYLKPLAGVYRSLKKQIEKNIFTF----NLNLNDILNSRLKKRKYFLDVLE---SDLM 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5.8%; Score 119; DB 3; Length 677; 22.2%; Pred. No. 0.37; Live 65; Mismatches 137; Indels 6
     -- LLNYKQEGDKCVENPNPTCNE 331
                               SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/295,802
FILING DATE: Herewith
                                                                                                                                                                                                                                                                  APPLICANT: Cochran Ph.D., Mark D
APPLICANT: Junker M.S., David E
ITILE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 188
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                   Sequence 115, Application US/08295802 Patent No. 6127163 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)977-9550
TELEX: 422523
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INDIVIDUAL ISOLATE: S-SPV-001
IMMEDIATE SOURCE:
CLONE: 515-85.1
                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE: John P. White STREET: 30 Rockefeller Plaza
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA (genomic)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: Swinepox virus
STRAIN: Kasza
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 677 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM:
296 SGCFRHLDEREECKC-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 76; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CITY: New York
STATE: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POSITION IN GENOME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGI.
MOLECULE TYPE: DN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HYPOTHETICAL: NAMTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MAP POSITION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORIGINAL SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ZIP: 10112
                                                                                                                                                            US-08-295-802-115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; UNITS: %G
US-08-295-802-115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
```

셤 ò Q ð

18;

```
Search completed: August 8, 2001, 12:29:56
Job time: 44 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   185 DAFHIPPTHISLRSL-RDINRIIELLKKYPNNNIIDYISDSIKSNSSFIHILHMIISNMF 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         244 PAIIPSVNDFISTVVDKDKLINMYGIKCVAMFSYDINMIDLESLDDSDYIFIEKNISIYD 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | :| |: | |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            993 SATIK---SKLLALRAVKNWKCYSLTNVSMYKKIKGVIVMDMVDYISTNILKYHKQLYDK 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           124 KVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNN 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          184 LVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLIN 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         244 DDTKKDMLGKLLSTGLVQNF----PNTIISKLIEGKF-QDMLNISQHQCVK--KQCPEN 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      71 QFKHISSNEYI--IEDSFKLLNSEQ-KNTLLKSYKY----IKESVENDIKFAQEGISYYE 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18 DVIYLKPLAGVYRSLKKQIEKNIFTF----NLNLNDILNSRLKKRKYFLDVLE---SDLM 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5.8%; Score 119; DB 4; Length 677;
ilarity 22.2%; Pred. No. 0.37;
Conservative 65; Mismatches 137; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,237A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 42A
ATTORNEY/ACTURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                450 MSTFEYKRDIKSCKCSICSDSITHHIYETTSCI-NYKSTDND 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   296 SGCFRHLDEREECKC-----LLNYKQEGDKCVENPNPTCNE 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David E.
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 115, Application US/08488237A Patent No. 6251403 GENERAL INFORMATION:
CURRENT FILING DATE: 1996-07-25
NUMBER OF SEQ ID NOS: 231
SOFTWARE: PALENLIN VEr. 2.1
SEQ ID NO 193
LENGTH: 677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISICS:
LENGTH: 677 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FELECOMMUNICATION INFORMATION
                                                                                                                                                                                                                                                                                         ) ORGANISM: Swinepox virus
US-08-686-968C-193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 76; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New York
: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-488-237A-115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STATE:
                                                                                                                                                                                                                                                      TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
```

```
347 ENNKVEEVLIHID-----NVSKNNK----LSLSDISSLMDQFRL----NPCTIRNILLS 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  393 SATIK---SKLLALRAVKNWKCYSLTNVSMYKKIKGVIVMDMVDYISTNILKYHKQLYDK 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         244 PAIIPSVNDFISTVVDKDRLINMYGIKCVAMFSYDINMIDLESLDDSDYIFIEKNISIYD 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 184 LVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLIN 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    244 DDTKKDMLGKLLSTGLVQNF-----PNTIISKLIEGKF-QDMLNISQHQCVK--KQCPEN 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         71 QFKHISSNEYI--IEDSFKLLNSEQ-KNTLLKSYKY----IKESVENDIKFAQEGISYYE 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              124 KVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNN 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18 DVIYLKPLAGVYRSLKKQIEKNIFTF----NLNLNDILNSRLKKRKYFLDVLE---SDLM 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; Score 119; DB 4; Length 677;
; Pred. No. 0.37;
65; Mismatches 137; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MSTFEYKRDIKSCKCSICSDSITHHIYETTSCI-NYKSTDND 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SGCFRHLDEREECKC-----LLNYKQEGDKCVENPNPTCNE 331
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 22.2%;
Matches 76; Conservative 65
                                                                                                                                                                                   : Swinepox virus
Kasza
                                                                                                                                                                                                                             INDIVIDUAL ISOLATE:
IMMEDIATE SOURCE:
CLONE: 515-85.1
                                                                                                                                                                                                                                                                                     CLONE: 515-85.1
POSITION IN GENOME:
MAP POSITION: ~2
                                                                                                                                                           ORIGINAL SOURCE
                                                                                                                                                                                                                                                                                                                                                                ; UNITS: %G
                                                                                                                                                                                   ORGANISM:
STRAIN: K
                                                                                                                                ANTI-SENSE:
```

18;

us-09-500-376-2.rai

PfMSP1(p19)S prote PfMSP1(p19)S prote MSP1EGF2A EGF2-11k MSP1EGF1B EGF1-11k MSP1EGF1B EGF1-11k MSP1EGF1A EGF1-11k

Plasmodium falcipa Plasmodium falcipa Plasmodium falcipa Plasmodium falcipa Plasmodium falcipa Merozite apical-en Merozoite apical-en Merozoite apical-en Merozoite apical-en Merozoite apical-en Plasmodium falcipa Plasmodium falcipa

Plasmodium falcipa H. pylori ORF 06gp plasmodium falcipa Plasmodium falcipa Plasmodium falcipa

þλ

Query Match

Score

Result No.

57.8 57.8 57.4 57.4 54.5 31.6 31.0

1283.5 1191.5 1191.5 1191.5 1191.5 1184.5 1184.5 651.5 640 640

10 9 10 11

Plasmodium falcipa Human integrin bet Integrin beta-1 ch

7

Perfect score:

Sequence:

OM protein

ö Run Scoring table:

Searched:

Minimum DB Maximum DB

Database

```
gp190; malaria; MSP-1; merozoite surface protein; stability; vaccine; monoclonal antibody; passive immunisation; parasite.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Recombinant production of complete gp190/MSP-1 Plasmodium surface protein - useful in anti-malaria vaccines, also stabilising genes reducing their AT content
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ALIGNMENTS
AAB37609
AAB37609
AAB32593
AAW325102
AAR41355
AAR41355
AAR41354
AAR41356
AAR41356
AAR41356
AAR41356
AAR41356
AAR1361821
AAB18216
AAB18205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P. falciparum synthetic gp190 protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAW54145 standard; Protein; 1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97WO-EP05441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        96DE-4040817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Tolle R;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1998-240088/21.
N-PSDB; AAV21451, AAV35363.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                108
96
127
127
54
54
49
49
49
49
3350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           H, Pan W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (BUJA/) BUJARD H.
  WO9814583-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              02-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        02-OCT-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23-SEP-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic.
    133
129.5
129.5
128
127.5
127.5
125.5
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.1
125.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                123
123
123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAW54145;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bujard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAW54145
    RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      P. falciparum synt
Merozoite surface
Merozoite surface
Merozoite surface
Merozoite surface
Modified merozoite
Sequence of the P1
Merozoite surface
PfMSP1(p19)A prote
PfMSP1(p19)A prote
                                                                                                                                                                     (without alignments)
393.312 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                 2062
1 AISVTMDNILSGFENEYDVI......SNFLGISFLLILMLILYSFI 394
                                                                                                                                             ; Search time 60.73 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | Sings | Acgdata | Geneseq | Jeneseq | Jenese
                     GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                     Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                            412676 segs, 60623988 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAW54145
AAY09372
AAY05832
AAY05833
AAY05833
AAY05834
AAY05834
AAY05834
AAW36103
AAW36103
AAW36103

    protein search, using sw model

                                                                                                                                             8, 2001, 12:29:12
                                                                                                                                                                                                                                                                                                              BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                            seq length: 0 seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A_Geneseq_0601:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              119
220
220
220
220
118
118
                                                                                                                                                                                                                            US-09-500-376-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length
                                                                                                                                             August
```

```
Sequence
                             Chen LH,
                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY05832;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1.27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EX PX PX PS
 셤
                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                      5
                           This sequence represents a modified Plasmodium falciparum gp190/MSP-1 (merozolte surface) protein. The gene encoding this protein has been stabilised by reducing the AT content of the nucleotide sequence. Such a protein is useful in vaccines against malaria or for producing monoclonal antibodies (for passive immunisation). The complete gp190 protein can now the produced outside the parasite and has, at least over extended regions, the native pattern of folding. Larger amounts of the protein can be produced recombinantly than would be possible using the parasites as
                                                                                                                                                                                                                1270 idnilskieneyevlylkplagvyrslkkglennvmtfnvnvkdilnsrfnkrenfknvl 1329
                                                                                                                                                                                                                                                                                                                                                       1489
                                                                                                                                                                                                                                                            1390 sekyksdldsikkyi-------ndkggenekylpflnnietlyktvn 1429
                                                                                                                                                                                                                                                                                                                                                                                      364
                                                                                                                                                                                                                                                 66 ESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                         126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                          246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDE 304
                                                                                                                                                                                     Gaps
                                                                                                                                                                                                          MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                                                                                  186 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD
                                                                                                                                                                                                                                                                                                                                              REECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSY
                                                                                                                                                                 DB 19; Length 1639;
                                                                                                                                                                                    21;
                                                                                                                                                                                    73; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            protein engineering; protein expression; codon usage;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MSP-1; merozoite surface protein; malaria; vaccine;
                                                                                                                                                                             7e-82;
                                                                                                                                                                ; Score 1283.5;
; Pred. No. 5.7e
54; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Merozoite surface protein MSP-1-42.
        Example 1; Fig 3c; 48pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY09372 standard; Protein; 355
                                                                                                                                                                 62.28;
62.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98US-0085649.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98WO-US22226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Plasmodium falciparum.
                                                                                                                                    1639 AA;
                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      transgenic animal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         31-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO9920774-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15-MAY-1998;
20-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20-OCT-1998;
                                                                                                                                                                                    242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29-APR-1999
                                                                                                                                    Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY09372;
                                                                                                                                                                            Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                   305
                                                                                                                                                                                                                                                                                                                                                                                                                                                           365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                  6
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
```

```
The present sequence represents a 42 kDa C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an importent target for the development of a vaccine against importent target for the development of a vaccine against the plasmodium falciparum. The nucleic acid sequence encoding MSP-1-42 has been modified (see AAX56008) compared to the native sequence (see AAX56009) such that 306 nucleotide positions have been replaced to lower the AT content (from 76 to 49.7%) and to eliminate 10 mRNA instability motifs while maintaining the same protein amino acid sequence. These alterations allow MSP-1-42 to be expressed in mammalian cell culture and in transgenic mice. Native MSP-1-12 is known to be difficult to express in cell culture systems. Commamalian cell culture systems or in transgenic animals. The invention allows expression of MSP-1 protein in the milk of transgenic animals, and also provides a DNA vaccine comprising a vector containing the altered MSP-1-42 sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ||| ||:|||| || sekyksdldsikkyi-------ndkqgenekylpflnnietlyktvn 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      287 reeckcllnykgegdkcvenpnptcnennggcdadakcteedsgsngkkitcectkpdsy 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        66 ESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDE 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 20; Length 355;
                                                                                                                                                                                                                            Modified malarial protein for use in anti-malarial vaccines
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  57.8%; Score 1191.5; DB 20; Lengue 60.2%; Pred. No. 2.4e-76; ... wiematches 72; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Merozoite surface protein MSP-1-42.
(GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY05832 standard; Protein; 355
                                                                                                                                                                                                                                                                                            Example; Fig 1; 35pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches 222; Conservative
                                                                                                                         WPI; 1999-288313/24
P-PSDB; AAX56008.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           347 plfdgifcs 355
                                                                ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     365 PLFDGIFCS 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               02-AUG-1999
```

2;

```
2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Important target for the development of a vaccine against
Important target for the development of a vaccine against
Inflation and falciparum. The nucleic acid sequence encoding MSP-1-42
Islamoduum falciparum. The nucleic acid sequence encoding MSP-1-42
Islamoduum falciparum. The nucleic acid sequence encoding MSP-1-42
Instability motifs while maintaining the same protein amino acid sequence. These alterations allow MSP-1-42 to be expressed in mammalian cell culture and in transgenic mice. The invention provides modified recombinant nucleic acid sequences and methods for increasing the mRNA levels and protein expression of proteins that are difficult to express in cell culture systems, mammalian cell culture systems or in transgenic animals. The preferred difficult corpusing such as parasites, bacteria and viruses that have DNA coding sequences of high AT content or which have mRNA instability motifs or rare codons relative to the recombinant expression system and machine and expression system.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    66 ESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New modified recombinant nucleic acid sequences useful for producing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   to be used. The invention allows expression of MSP-1 protein in the milk of transgenic animals, and also provides a DNA vaccine comprising a vector containing the altered MSP-1-42 sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 20; Length 355;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       present sequence represents a 42 kDa C-terminal portion of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           57.8%; Score 1191.5; DB 20; Length
60.2%; Pred. No. 2.4e-76;
.ive 54; Mismatches 72; Indels
                                  engineering; protein expression; codon usage;
                MSP-1; merozoite surface protein; malaria; vaccine;
                                                                                                                                                                                                                                                                         (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Disclosure; Fig 1; 43pp; English
                                                                                                                                                                                      98WO-US22225.
                                                                                                                                                                                                                       98US-0085649.
97US-0062592.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity 60.2
wes 222; Conservative
                                                                                 Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                   malarial DNA vaccine
                                                                                                                                                                                                                                                                                                                                          WPI; 1999-302742/25.
N-PSDB; AAX25586.
                                                                                                                                                                                                                                                                                                            Chen LH, Meade H;
                                                  transgenic animal.
                                                                                                                   W09920766-A2
                                                                                                                                                                                    20-OCT-1998;
                                                                                                                                                                                                                        15-MAY-1998;
                                                                                                                                                                                                                                         20-OCT-1997;
                                                                                                                                                    29-APR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sedneuce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
```

```
246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDE 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This present sequence comprises a 42 kDa C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against plasmodium falciparum. The C-terminal end of the sequence is modified to include a 6xHis tag. A nucleic acid (see AAX5608) encoding MSP-1-42 has been modified according to a method of the invention in order to improve expression in mammalian cells and in transgenic animals by reducing the AT content and removing mRNA instability motifs. The invention allows expression of MSP-1-42 in the milk of transgenic animals, and also provides a DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   vaccine comprising a vector containing the altered MSP-1-42 nucleic
             REECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 20; Length 361;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  protein for use in anti-malarial vaccines
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                       MSP-1; merozoite surface protein; malaria; vaccine; protein engineering; protein expression; codon usage;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    57.8%; Score 1191.5; DB llarity 60.2%; Pred. No. 2.5e-76; Conservative 54; Mismatches 72
                                                                                                                                                                                                                                                                                            Merozoite surface protein MSP-1-42.
                                                                                                                                                                                                         AAY09373 standard; Protein; 361 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Example; Fig 2; 35pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98US-0085649.
97US-0062592.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US22226.
                                                                                                                                                                                                                                                                                                                                                                              Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-288313/24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Modified malarial
                                                                                                                            347 plfdgifcs 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                          PLFDGIFCS 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Meade H;
                                                                                                                                                                                                                                                                                                                                                   transgenic animal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              361 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P-PSDB; AAX56009
                                                                                                                                                                                                                                                                                                                                                                                                         W09920774-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                20-0CT-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20-OCT-1997;
                                                                                                                                                                                                                                                               31-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                    29-APR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Simes 222;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local S
Matches 222
                                                                                                                                                                                                                                     AAY09373;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Chen LH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                          365
                                                    305
                                                                                                                                                                                            AAY09373
                       g
                                                                                                         δý
                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                    ŏ
```

ä

```
W09920774-A2
 88888888888
                                                                                                                                                                                   g
                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                           ολ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                    δ
           186 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD 245
ESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                      246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDE 304
                                                                                                                                                               REECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSY 364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This present sequence comprises a 42 kDa C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against Plasmodium falciparum. The C-terminal end of the sequence is modified to include a 6xHis tag. Nucleic acids (see AAX25586 and AAX25593) encoding MSP-1-42 have been modified according to a method of the invention in order to improve expression in mammalian cells and in transgenic animals. The invention provides modified recombinant nucleic acid sequences and methods for increasing the mRNA levels and protein expression of proteins that are difficult to express in cell culture systems, especially mammalian cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   modified recombinant nucleic acid sequences useful for producing
                                      126 LAKYKODLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEOKKESKFLPFLTNIETLYNNLV
                                                                                                                                  MSP-1; merozoite surface protein; malaria; vaccine; protein engineering; protein expression; codon usage;
                                                                                                                                                                                                                                                                                                                                      Merozoite surface protein MSP-1-42.
                                                                                                                                                                                                                                                                           AAY05833 standard; Protein; 361 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Disclosure; Fig 2; 43pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            98WO-US22225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98US-0085649.
97US-0062592.
                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                               Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            malarial DNA vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1999-302742/25.
                                                                                                                                                                                                     PLFDGIFCS 373
                                                                                                                                                                                                                     347 plfdgifcs 355
                                                                                                                                                                                                                                                                                                                                                                             transgenic animal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ή
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Chen LH, Meade
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAX25587
                                                                                                                                                                                                                                                                                                                                                                                                                    W09920766-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                           20-OCT-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15-MAY-1998;
20-OCT-1997;
                                                                                                                                                                                                                                                                                                                   02-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                        29-APR-1999
                                                                                                                                                                                                                                                                                              AAY05833;
                                                                                                                                                              305
99
                                                                                                                                                                                                      365
                                                                                                                                                                                                                                                                g
                                                                                                                                                                                 g
                                                                                                                                                                                                                       8
ð
                9
                                       ò
                                                       9
                                                                             ò
                                                                                                                     ò
                                                                                                                                        g
                                                                                                                                                              ò
                                                                                                                                                                                                      ò
```

```
245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDE 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        305 REECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSY 364
                                                                                                                                                                                                                                                                                                                                        21; Gaps
                                                                                                                                                                                                                                                                                                                                                                                    6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVL 65
culture systems or in transgenic animals. The preferred difficult protein candidates for expression are those derived from lower organisms such as parasites, bacteria and viruses that have DNA coding sequences of high AT content or which have mRNA instability motifs or rare codons relative to the recombinant expression system to be used. The invention allows expression of MSP-1 in the milk of transgenic animals, and also provides a DNA vaccine comprising a vector containing the altered MSP-1-42 nucleic acid.
                                                                                                                                                                                                                                                                                                                                                                                                                 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note= "Asn in native MSP-1-42 (N181Q mutation)"
278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note= "Asn in native MSP-1-42 (N262Q mutation)"
                                                                                                                                                                                                                                                                                          DB 20; Length 361;
                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MSP-1; merozoite surface protein; malaria; vaccine; protein engineering; protein expression; codon usage; transgenic animal; mutant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1..15
/note= "beta-casein signal peptide"
                                                                                                                                                                                                                                                                                       57.8%; Score 1191.5; DB 20
60.2%; Pred. No. 2.5e-76;
ive 54; Mismatches 72;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Modified merozoite surface protein MSP-1-42.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY09374 standard; Protein; 376 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "6xHis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    371..376
/note= "6
                                                                                                                                                                                                                                                                                                           Best_Local Similarity 60.2
Matches 222; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             347 plfdgifcs 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PLFDGIFCS 373
                                                                                                                                                                                                                      361 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             31-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Synthetic.
                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY09374;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY09374
```

```
New modified recombinant nucleic acid sequences useful for producing
                                                                                                                                        /note= "Asn in native MSP-1-42 (N181Q mutation)"
278
                                                                                                                                                                    /note= "Asn in native MSP-1-42 (N262Q mutation)"
                     protein engineering; protein expression; codon usage;
transgenic animal; mutant.
             MSP-1; merozoite surface protein; malaria; vaccine;
                                                                                             1..15
∕note≈ "beta-casein signal peptide"
                                                                                   Location/Qualifiers
                                                                                                                                                                                                                                                                                       (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                             Example; Fig 11; 43pp; English.
                                                                                                                            /note= "6xHis
                                                                                                                                                                                                                                                        98US-0085649.
                                                                                                                                                                                                                                    98WO-US22225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 221; Conservative
                                                    Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                         malarial DNA vaccine
                                                                                                                                                                                                                                                                                                                              1999-302742/25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                             Meade H;
                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAX25593
                                                                                                                                      Misc-difference
                                                                                                                                                          Misc-difference
                                                                                                                                                                                        WO9920766-A2
                                                                                                                                                                                                                                   20-0CT-1998;
                                                                                                                                                                                                                                                       15-MAY-1998;
20-OCT-1997;
                                                                                                                                                                                                               29-APR-1999
                                                               Synthetic
                                                                                                                                                                                                                                                                                                            Chen LH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                    protein
                                                                                             Peptide
                                                                                                                  Peptide
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                    7
                                                                                                                                                                                                                                                                                                                                                                                                                            66 ESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                        126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     186 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDE 304
                                                                                                                                                                                                                                                         t
                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                       6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                                                                                                                                  The present sequence represents a 42 kDa C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against Plasmodium falciparum. The sequence has been modified to include an N-terminal beta-casein signal peptide and an N-terminal of SKH is and N2620 mutations have been introduced eliminate 2 N-01ycosylation sites. These modifications allow the MSP-1-42 protein to be expressed in the milk of transgenic mice.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              242 nhnnlltkflstgmvfenlaktvlsnlldgnlqgmlqisqhqcvkkqcpqnsgcfrhlde
                                                                                                                                                                                                                                                                                                                                                  21;
                                                                                                                                                                                                                                                                                                                               Length 376;
                                                                                                                                                          protein for use in anti-malarial vaccines
                                                                                                                                                                                                                                                                                                                             ; Score 1184.5; DB 20; Lengt; Pred. No. 8.1e-76; 53; Mismatches 74; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Modified merozoite surface protein MSP-1-42.
                                                                                 (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                              Example; Fig 11; 35pp; English
                                                                                                                                                                                                                                                                                                                              57.4%;
                                                  98US-0085649.
                              98WO-US22226
                                                                                                                                                                                                                                                                                                                                                  Matches 221; Conservative
                                                                                                                         WPI; 1999-288313/24.
P-PSDB; AAX56008.
                                                                                                                                                                                                                                                                                                                                        Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY05834 standard;
                                                                                                                                                        Modified malarial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         365 PLFDGIFCS 373
                                                                                                      Chen LH, Meade
                             20-OCT-1998;
                                                  15-MAY-1998;
20-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              02-AUG-1999
         29-APR-1999
                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY05834;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAYO5834
ID AAYO
XX
AC AAYO
XX
XX
DT 02-A
XX
DE MOdi
g
                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
The present sequence represents a 42 kDa C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against plasmodium falciparum. The sequence has been modified to include now terminal beta-casein signal peptide and an N-terminal 6xHis tag. In addition, N1810 and N2620 mutations have been introduced to eliminate 2 N-glycosylation sites. These modifications allow the MSP-1-42 protein to be expressed in the milk of transgenic mice. The invention generally relates to modified recombinant nucleic acid sequences and methods for increasing the mRNA levels and protein expression of proteins that are difficult to express in call culture systems. The preferred difficult protein candidates for transgenic animals. The preferred difficult protein candidates for bacteria and viruses that have DNA coding sequences of high AT content or which have mRNA instability motifs or rare codons relative to the recombinant expression system to be used.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ಧ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               66 ESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 20; Length 376;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              74; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  57.4%; Score 1184.5; DB 59.9%; Pred. No. 8.1e-76;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   53; Mismatches
```

245

ဖ

```
1284 idnilskieneyevlylkplagvyrslkkglennvmtfnvnvkdilnsrfnkrenfknvl 1343
                         ESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                  EREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDS 363
                                                                                                                                                                                               LAKYKDDLESIKKVIKEEKEKPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV
                                                                                                                              NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD
                                                                                                                                                                                246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKF-QDMLNISQHQCVKKQCPENSGCFRHLD
                                                                                                                                                                                                                                                                                    364 YPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                    1404 sekyksdldsikkyi-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Merozoite
                                                                                                                                                                                                                                                                                                                                                                                                      AAB37610;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        malaria
                                                                                                                                                                                                                                                                                                                                                    σ
                                                                            126
                         99
                                                                                                                               186
                                                                                                                                                                                                                                    304
                                                                                                                                                                                                                                                                                                                                                                AAB37610
                                                  g
                                                                                                                                                      g
                                                                                                                                                                                                       g
                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                   ò
                        οy
                                                                                                                              οy
|:|: |:||| |||||| :
--ndkggenekylpflnnietlyktvn 181
                                                                                                                                                         The sequence encoding the P195 protein of Plasmodium falciparum (AAN50530) and a peptide comprising at least one of its epitopes (see AAP50777) are claimed. Also claimed is a vaccine for inducing immunity to malaria comprising the novel peptide or P195 or a peptide comprising at least one epitope when derived from the new DNA sequence, together with a carrier.
                                                    NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD
                                                                                         246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDE
                                                                                                                                           305 REECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sequence encoding plasmodium falciparum protein - expressing the protein for use in vaccines against
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Riveros-Moreno V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 1654;
                                                                                                                                                                                                                                                                                                                                                                Sequence of the P195 protein of Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             54.5%; Score 1124; DB 6;
56.3%; Pred. No. 1e-70;
ive 60; Mismatches 89;
                                                                                                                                                                                                                                                                                                                                                                                            vaccine; epitope; antigen; immunogen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Lockyer M,
                                                                                                                                                                                                                                                                                    AAP50777 standard; Protein; 1654 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (WELL ) WELLCOME FOUNDATION LTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    χ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 6; Fig 1; 51pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      84GB-0024340.
84GB-0004692.
85GB-0004429.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Odink
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               85EP-0301173
                                                                                                                                                                                                                                                                                                                                        (first entry)
            142 sekyksdldsikkyi--
                                                                                                                                                                                                                                                                                                                                                                                                                   Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sandhu J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 1985-224845/37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1654 AA;
                                                                                                                                                                                                            362 plfdgifcs 370
                                                                                                                                                                                              365 PLFDGIFCS 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; AAN50530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26-SEP-1984;
22-FEB-1984;
21-FEB-1985;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21-FEB-1985;
                                                                                                                                                                                                                                                                                                                                        30-SEP-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11-SEP-1985
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Cloned DNA useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                            EP154454-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Holder A,
                                                                                                                                                                                                                                                                                                              AAP50777;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                          Malaria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              malaria
                                                                                                                                                                                                                                                                                                  g
                                                                                                            a
                                                                                                                                           ð
                                                                                                                                                                    a
                                                                                                                                                                                            ò
                                                                                                                                                                                                                   g
                                      ò
                                                                                      ò
```

```
The present invention relates to non-natural variants of a C-terminal fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The non-natural variants have reduced affinity for at least 1 antibody capable of blocking a second antibody that inhibits the proteolytic cleavage of Plasmodium MSP-1_4_2, and has the same affinity for at least one third antibody that inhibits the proteolytic cleavage of plasmodium MSP-1_4_2, and has the same affinity for at least MSP-1_4_2, compared to natural MSP-1_1_9. The non-natural variants of the present invention are useful for immunising a mammal against malaria, and can be used to treat malaria. The present sequence is MSP-133 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ပ်
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Uthaipibull
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            treating or preventing
                                                                                                                                                                                                             protein; protazoacide; vaccine; malaria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Novel variants of the C-terminal fragment of Plasmodium surface protein-1, useful as vaccines for treating or pr
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Syed S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                    3
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Morgan
A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     B, Feeney J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Example 5; Fig 15; 126pp; English.
                                                                                       AAB37610 standard; Protein; 262
                                                                                                                                                                                Merozoite surface protein-133.
                                                                                                                                                                                                                                                                                                                                                            99GB-0009072.
99US-0311817.
99CA-2271451.
                                                                                                                                                                                                                                                                                                                                                                                                                      (MEDI-) MEDICAL RES COUNCIL
                                                                                                                                                                                                                                                                                                                                  20-APR-2000; 2000WO-GB01558
                                                                                                                                                                                                                                         Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Holder A, Birdsall
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2001-015762/02.
                                                                                                                                                                                                               surface
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-PSDB; AAC68978
                                                                                                                                                                                                                                                                       WO200063245-A2
                                                                                                                                                                                                                                                                                                                                                              20-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                         13-MAY-1999;
25-MAY-1999;
                                                                                                                                                   27-FEB-2001
                                                                                                                                                                                                                                                                                                    26-0CT-2000
```

٠. ن

Gaps

22;

Indels

Local Similarity 56.3 Hes 220; Conservative

Best Loca Matches

õ

Query Match

6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVL 65

ö

us-09-500-376-2.rag

XX So

g ò a g

ò

셤 ò g

δ

```
Disclosure; Fig 1B; 85pp; French
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97WO-FR00291.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                96FR-0001821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PASTEUR.
NEW YORK STATE.
                                                                                                                                                                                                                                                                                                                                                                                                                                        Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1997-425034/39.
                                                                                                                     116 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P-PSDB, AAW22592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (INSP ) INST
(UYNY ) UNIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO9730159-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            14-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                14-FEB-1996;
                                                                                                                                                                                                                                                                                                                                                              25-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21-AUG-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Synthetic.
                                                                                                                                                                                                                                                                                                                                         AAW22592;
                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Roth C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Region
                                                                                                                                                                                                                                                                                               11
 8 × 0 0 0 0 0 0 0 0 0 × 8
                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                            g
                                                              ς;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    96..116
/note= "glycosylphosphatidylinositol anchoring sequence"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note= "amino acids derived from P. falciparum MSP1 p19 fragment"
                                                                                                                                         LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                             ESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                             Recombinant protein containing the merozoite surface protein-1 p19 fragment - useful in anti-malarial vaccines, diagnosis and protein purification
                                                                                   MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVL
                                                                                                                                                                                                                NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD
                                                              21;
                                          Length 262;
                                                                                                                                                                                                                                                                                                                                                                                                                          Plasmodium vivax; merozoite surface protein; MSP1; p19;
Plasmodium falciparum; malaria; vaccine; immunity; epitope.
                                       31.6%; Score 651.5; DB 22; Length 47.1%; Pred. No. 1.9e-38; ive 52; Mismatches 73; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Nato
                                                                                                                                                                                                                                                                         227 nhnnlltkflstqmvfenlaktvlsnlldqnlqqml 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mendis K,
                                                                                                                                                                                                                                                             246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDML
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                      AAW36103 standard; Protein; 116 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Barnwell JW, Longacre-Andre S,
                                                                                                                                                                                                                                                                                                                                                                                                    PfMSP1(p19)A protein sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               97WO-FR00290.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    96FR-0001822.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (INSP ) INST PASTEUR.
(UYNY ) UNIV NEW YORK STATE.
                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                  Best Local Similarity 47.1
Matches 130; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1997-425033/39
          Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAT94550
         262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO9730158-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               14-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                 25-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   14-FEB-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-AUG-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic.
                                                                                                                                                                                                                                                                                                                                                          AAW36103;
                                         Query Match
          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Key
Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Roth C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Region
                                                                                                                             99
                                                                                                                                                 29
                                                                                                                                                                                                                 186
                                                                                                                                                                                                                                                                                                                 RESULT 1
AAW36103
```

```
This is the amino acid sequence of a recombinant protein comprising amino acids 1613-1705 of the Plasmodium falciparum merozoite surface protein 1 (MSP1) 19 KD C-terminal fragment (p19), linked to a glycosylphosphatidylinositol membrane anchoring sequence. p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the p19 fragment provides a high lavel of protective immunity since it includes epitopes not presented in the p42 fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       96..116
/note= "glycosylphosphatidylinositol anchoring sequence"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             281 NISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADA 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "amino acids derived from P. falciparum MSP1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Recombinant protein containing Plasmodium merozoite surface protein-1 p42 fragment - useful in antimalarial vaccines, also new antibodies for diagnosis and protein purification
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    341 KCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                     Length 116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Plasmodium vivax; merozoite surface protein; MSP1; p19;
Plasmodium falciparum; malaria; vaccine; immunity; epitope.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ..
[24
                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match
31.0%; Score 640; DB 18;
Best Local Similarity 100.0%; Pred. No. 4.2e-38;
Matches 114; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Nato
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mendis K,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAW22592 standard; Protein; 116 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Disclosure; Fig 1B; 85pp; French
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Barnwell JW, Longacre-Andre S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PfMSP1(p19)A protein sequence
```

œ

```
The recombinant protein can be used for the production of anti-malarial vaccines, where the p19 fragment provides a high level of protective immunity since it includes epitopes not presented in the p42 fragment.
                                                                                        p19 is the
This is the amino acid sequence of a recombinant protein comprising amino acids 1613-1705 of the Plasmodium falciparum merozoite surface protein 1 (MSP1) 19 kD C-terminal fragment (pll), inlked to a glycosylphosphatidylinositol membrane anchoring sequence. pl9 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species.
     888888888888
```

116 AA; Sequence

```
;
0
                                          281 NISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADA 340
                        Gaps
                                                  ;
0
      Length 116;
                        Indels
   Score 640; DB 18;
Pred. No. 4.2e-38;
0; Mismatches 0;
31.0%; Scor.
100.0%; Pred. No. 4...
   Query Match
Best Local Similarity 100.
Matches 114; Conservative
                                                          셤
                                          ò
                                                                             ŏ
                                                                                             g
```

Merozoite surface protein; protazoacide; vaccine; malaria. Ź AAB37609 standard; Protein; 108 Merozoite surface protein-119. (first entry) 27-FEB-2001 AAB37609; RESULT 12 AAB37609

99GB-0009072. 99CA-2271451. 20-APR-2000; 2000WO-GB01558 Plasmodium falciparum WO200063245-A2 20-APR-1999; 13-MAY-1999; 25-MAY-1999; 26-OCT-2000 

J, Morgan W, Feeney B, Birdsall WPI; 2001-015762/02 N-PSDB; AAC68977 Ä Holder

(MEDI-) MEDICAL RES COUNCIL

ပ်

Uthaipibull

s, Syed

Example 2; Page 48; 125pp; English.

Novel variants of the C-terminal fragment of Plasmodium merozoite surface protein-1, useful as vaccines for treating or preventing malaria -

Example 5; Fig 15; 126pp; English.

fragment of a plasmodium merozoite surface protein-1 (MSP-1). The non-natural variants have reduced affinity for at.least 1 antibody capable of blocking a second antibody that inhibits the proteolytic cleavage of plasmodium MSP-1\_4\_2, and has the same affinity for at least one third antibody that inhibits the proteolytic cleavage of Plasmodium MSP-1\_4\_2, and proteolytic cleavage of Plasmodium MSP-1\_4\_2, more an unitarial variants of the present invention are useful for immunising a mammal against malaria, and can be used to treat malaria. The present sequence is MSP-119 protein. present invention relates to.non-natural variants of a C-terminal The

Ë 108 Sequence

```
ij
                                    272 IEGKFQDMLNISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCVENPNPTCNE 331
                                                                                                                                                                                                                                                                                                                                                                 ပဲ
                                                       8 iegrh----niaghqcvkkqcpqnsgcfrhldereeckcllnykqegdkcvenpnptcne 63
                                                                                                                                                                                                                                                                                                                                                                 Uthaipibull
                                                                                                                                                                                                                                                                                                                                                                                                   Novel variants of the C-terminal fragment of Plasmodium merozoite surface protein-1, useful as vaccines for treating or preventing
                    4;
Length 108;
                                                                                                                                                                                                                 Merozoite surface protein; protazoacide; vaccine; malaria
                   Indels
                                                                         NNGGCDADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSN 376
                                                                                  ŝ
                                                                                                                                                                                                                                                                                                                                                                 Syed
Score 554; DB 22;
Pred. No. 4.5e-32;
                                                                                                                                                                                                                                                                                                                                                                 Morgan W,
                  3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                 J,
                                                                                                                                         AA.
                                                                                                                                       AAB37608 standard; protein; 96
                                                                                                                                                                                                                                                                                                                                                                 Feeney
26.9%;
92.4%;
                                                                                                                                                                                                                                                                                                                   99US-0311817.
99CA-2271451.
                                                                                                                                                                                              Merozoite surface protein-1.
                                                                                                                                                                                                                                                                                                            99GB-0009072.
                                                                                                                                                                                                                                                                                                                                                (MEDI-) MEDICAL RES COUNCIL.
                                                                                                                                                                                                                                                                                         20-APR-2000; 2000WO-GB01558
                  Conservative
                                                                                                                                                                                                                                                                                                                                                                 Holder A, Birdsall B,
                                                                                                                                                                                                                                   Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2001-015762/02.
         Similarity
                                                                                                                                                                                                                                                     WO200063245-A2
                                                                                                                                                                                                                                                                                                            20-APR-1999;
                                                                                                                                                                                                                                                                                                                   13-MAY-1999;
25-MAY-1999;
Query Match
Best Local Simi
Matches 97;
                                                                                                                                                                            27-FEB-2001
                                                                                                                                                                                                                                                                        26-OCT-2000
                                                                                                                                                          AAB37608;
                                                                                                                                                                                                                                                                                                                                                                                                                        malaria
                                                                                                                       13
                                                                         332
                                                                                          64
                                                                                                                               AAB37608
                                                                                                                                                          ò
                                                                                          g
```

ô The present invention relates to non-natural variants of a C-terminal fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The non-natural variants have reduced affinity for at least 1 antibody capable of blocking a second antibody that inhibits the proteolytic cleavage of plasmodium MSP-14.2, and has the same affinity for at least one third antibody that inhibits the proteolytic cleavage of Plasmodium MSP-12.9. The present sequence is the wild-type MSP-1 protein. This sequence was used to generate the variants of the present invention. The non-natural variants of the present invention are useful for immunishing a mammal against malaria, and can be 281 NISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADA 340 Gaps 1. nisqhqcvkkqcpqnsgcfrhldereeckcllnykqegdkcvenpnptcnennggcdada . 0 Length 96; Indels Score 553; DB 22; Pred. No. 4.5e-32; 1; Mismatches 0; 1; Mismatches 26.8%; 95; Conservative used to treat malaria. Local Similarity AA; 96 Sequence Query Match Matches ŏ q

Ø

ö

0; Gaps

Indels

.; ;;

Mismatches

ò g

```
ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (INSP ) INST PASTEUR.
(UYNY ) UNIV NEW YORK STATE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97WO-FR00290,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 96FR-0001822.
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      20..127
/note= "r
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1997-425033/39.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-PSDB; AAT94549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO9730158-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         14-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   14-FEB-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               purification
94;
                                                                                                                                                                                                                                                                                                                                        25-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21-AUG-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Chimeric
Chimeric
                                                                                                                                                                                                                                                                                                   AAW36102;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Key
Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Roth C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Region
Matches
                                                                                                                                                                                                                                          AAW36102
                                                                                                                                                                                                                       RESULT
                                                                           g
                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This is the amino acid sequence of a chimeric protein comprising amino acids 1-32 of the plasmodium vivax merozoite surface protein 1 (MSP1) linked to amino acids Asn1613-Ser1705 which correspond to the 19 kD c-terminal fragment (p19) of MSP-1 from Plasmodium falciparum. p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the p19 fragment provides a high level of protective immunity since it includes epitopes not presented in the p42 fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               33..34
/note= "encoded by restriction enzyme sequence used
  create the chimeric sequence"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Recombinant protein containing Plasmodium merozoite surface protein-1 p42 fragment - useful in antimalarial vaccines, also new antibodies for diagnosis and protein purification
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   35.127
/note- "derived from P. falciparum C-terminal p19
fragment of MSP1"
                                                                                                                                                                                                                                                                                                               Chimeric; Plasmodium vivax; merozoite surface protein; MSP1; p19; Plasmodium falciparum; malaria; vaccine; immunity; epitope.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /note= "derived from P. vivax MSP1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mendis K, Nato F;
                341 KCTEEDSGSNGKKITCECTKPDSYPLFDGIFCSSSN 376
                                   /note= "signal peptide"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note= "mature protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                      AAW22593 standard; Protein; 127 AA

    Plasmodium vivax.
    Plasmodium falciparum.

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Disclosure; Fig 1C; 85pp; French
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Barnwell JW, Longacre-Andre S,
                                                                                                                                                                                                                                                                            PfMSP1(p19)S protein sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (INSP ) INST PASTEUR.
(UYNY ) UNIV NEW YORK STATE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     97WO-FR00291.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            96FR-0001821
                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20..127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1997-425034/39.
P-PSDB; AAW22592.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              127 AA;
                                                                                                                                                                                                                                      25-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO9730159-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     14-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            14-FEB-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21-AUG-1997.
                                                                                                                                                                                                                                                                                                                                                                          Chimeric
                                                                                                                                                                                              AAW22593;
                                                                                                                                                                                                                                                                                                                                                                                               Chimeric
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Roth C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Region
                                                                                                                                                                                                                                                                                                                                                                                                                                         Key
                                                                                                               RESULT
```

Score 546; DB 18; Length 127; Pred. No. 2e-31;

26.5%; 97.9%;

Best Local Similarity

Query Match

```
acids 1-32 of the Plasmodium vivax merozoite surface protein 1 (MSP1) linked to amino acids Asn1613-Ser1705 which correspond to the 19 kD C-terminal fragment (p19) of MSP-1 from Plasmodium falciparum. p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the p19 fragment provides a high level of protective immunity since it includes epitopes not presented in the p42 fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This is the amino acid sequence of a chimeric protein comprising amino
278 DMLNISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCD 337
                /note= "encoded by restriction enzyme sequence used create the chimeric sequence"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Recombinant protein containing the merozoite surface protein-1 p19 fragment - useful in anti-malarial vaccines, diagnosis and protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            p19
                                                                                                                                                                                                                                                                  Chimeric; Plasmodium vivax; merozoite surface protein; MSP1; p19; Plasmodium falciparum; malaria; vaccine; immunity; epitope.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /note= "derived from P. falciparum C-terminal
fragment of MSP1"
                                                                                                                                                                                                                                                                                                                                                                                                                                       "derived from P. vivax MSP1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Nato F;
                                                       338 ADAKCTEEDSGSNGKKITCECTKPDSYPLFDGIFCS 373
                                                                      92 adakcteedsgsngkkitcectkpdsyplfdgifcs 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mendis K,
                                                                                                                                                                                                                                                                                                                                                                   1..19
/note= "signal peptide"
                                                                                                                                                                                                                                                                                                                                                                                                           "mature protein"
                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                                                                                                                                        Z
                                                                                                                                                                                                                                                                                                           Plasmodium vivax.
Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Fig 1C; 85pp; French.
                                                                                                                                                    AAW36102 standard; Protein; 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Barnwell JW, Longacre-Andre S,
                                                                                                                                                                                                                                        PfMSP1(p19)S protein sequence.
```

```
SQ Sequence 127 AA;
```

; 0 Gaps ö Query Match 26.5%; Score 546; DB 18; Length 127; Best Local Similarity 97.9%; Pred. No. 2e-31; Matches 94; Conservative 0; Mismatches 2; Indels (

QQ ò q

Search completed: August 8, 2001, 12:31:53 Job time: 161 sec

THIS PAGE BLANK (USTTO)

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

August 8, 2001, 12:29:12; Search time 43.63 Seconds (without alignments) 687.893 Million cell updates/sec Run on:

Title: Perfect score: Sequence:

US-09-500-376-2 2062 1 AISVTMDNILSGFENEYDVI......SNFLGISFLLILMLILYSFI 394

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

219241 seqs, 76174552 residues Searched:

219241 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

PIR\_68:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

SUMMARIES	B ID Description	1 CASOOM	major merozoite eurfac	A45948	major merozoite	A26868 major merozoite	A45545	805603		A24594			A28121 major				hypothetical	2 T18372 repeat organellar	A64505	T28676		T34081 hypothetical		B70356		S16752 major merozoi		E81339 probable rest	2 T18507 hypothetical prote
Σ <b>λ</b>	Length DB ID	1726 1 9220	10		7	N	7	7	_	7	7	7	7	7	7	7	7		7	7		7	7	7	7	86 2	7		~
œ	Query Match Le	0 001	1 66	99.1	98.3	97.3		62.	56.	54.	40.	40	29	29.8	28.3	8.8	8.3	7.9	7.7	7.6	7.5	7.4	_	7	7.0	7.0	•	6.9	8
	Score	2062	2043 5	2043	2027	2006	1289.5	1289.5	1166	1124	843	832	613.5	613.5	583	180.5	172	162.5	159.5	156	155.5	153	152.5	150	14	144.5	143	142	141
		-	10	'n	4	5	9	7	ω	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28

ORF MSV156 hypothe hypothetical prote	surface membrane p hypothetical prote chromosome segrega	hypothetical prote	nypotnetical prote P115 protein homol myosin heavy chain	reticulocyte-bindi hypothetical prote	nypotnetical prote hypothetical prote conserved hypothet	multimerin, endoth
55	5 5 3 5 4	2.9.5	55.55	123	37.5	4
T28317 T18355	PC600 H6450 A4946	C7162	F966/3 E70105 A26655	S4838	S67087	A5738
000	000	000	7 77 77	000	900	10
1127	624 671 1225	1979	1313 819 2116	2829	821 821 880	1228
6.8		6.7	6.6 6.6	9.9	900	6.5
141	140 139.5 139.5	138.5	137.5 137.5 137	136.5	135 135	135
30	33 34 34	3 22	38 38 36	440	4 4 4 7 6 4	45

## ALIGNMENTS

æ

ö

g

ò 셤 ò qq ò 8

```
C;Accession: A54498
R;Peterson, M.G.; Coppel, R.L.; McIntyre, P.; Langford, C.J.; Woodrow, G.; Brown, G.V
Mol. Biochem. Parasitlol. 27, 291-302, 1988
A;Title: Variation in the precursor to the major merozoite surface antigens of Plasmo A;Reference number: A54498; MUID:88142999
A;Scatus: preliminary
A;Scatus: preliminary
A;Residues: 1-1701 < PET>
A;Residues: 1-1701 < PET>
A;Csuperfamily: major merozoite surface antigen
C;Superfamily: major merozoite surface antigen
C;Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)
                 Exp. Parasitol. 67, 1-11, 1988
A;Title: Plasmodium falciparum: gene structure and hydropathy profile of the major A;Reference number: A45948; MUID:89003525
A;Reference number: A45948; MUID:89003525
A;Reference number: BNA
A;Reference number: BNA
A;Reference number: BNA
A;Reference number: BNA
C;Residues: 1-1726 <CHA>
C;Residues: 1-1726 <CHA>
A;Cross-references: GB:MA7213
C;Superfamily: major merozoite surface antigen
C;Reywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1333 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNINDILNSRLKKRKY 1392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C;Species: Plasmodium falciparum
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1633 HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 1692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1308 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY 1367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEOKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1393 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AISVIMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 AISVIMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HLDEREECKCLLNYKOEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                Length 1726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                              3;
                                                                                                                                                                                                                                                                                                                                                2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 2027; DB 2;
Pred. No. 1.4e-98;
                                                                                                                                                                                                                                                                                                                                              Score 2043; DB Pred. No. 2e-99; 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1693 PDSYPLEDGIFCSSSNFLGISFLLILMLILYSFI 1726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 98.3%; Score 2027; Dest Local Similarity 98.7%; Pred. No. 1.4e Matches 389; Conservative 1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI
                                                                                                                                                                                                                                                                                                                                              Query Match 99.1%;
Best Local Similarity 99.2%;
Matches 391; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Н
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    merc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (st
                                                                                                                                                                                                                                                                                                                                                                                                                           RO-71)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C-terminal region of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;;
                                                                                                                                                                                                                                                                                                                                                                                                                           parasite (Plasmodium falciparum) (strain
                                                                                                                                                                                                                                                                                                                                                                                                                merozoite surface antigen 1 - malaria parasite (Plasmodium falciparum) (stra C;Species: Plasmodium falciparum A;Variety: strain R0-71
A;Variety: strain R0-71
C;Species: D6-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 09-Jun-2000 C;Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 09-Jun-2000 C;Accession: 847282
R;Tolle, Rx.; Bujard, H.; Cooper, J.A.
A;Description: Plasmodium falciparum: recombination within the C-terminal re A;Reference number: 847282
A;Accession: 847282
A;Accession: 847282
A;Accession: 1-651 <TOL>
                                              major merozoite surrace ancasum falciparum
C;Species: Plasmodium falciparum
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jun-2000
1513 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK 1572
                                                                                                                                                                            1633 HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 1692
                                                                                                                                               HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Cross-references: EMBL:23529; NID:9535257; PIDN:CAA84558.1; PID:9535258
A;Experimental source: strain R0-71
C;Superfamily: major merozoite surface antigen
C;Keywords: glycoprotein; merozoite; surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99.1%; Score 2043.5; DB 2
ilarity 99.5%; Pred. No. 6.6e-100;
Conservative 1; Mismatches 0;
                                                                                                                                                                                                                                                                        1693 PDSYPLFDGIFCSSNFLGISFLLILMLILYSFI 1726
                                                                                                                                                                                                                                                 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 392; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    361
                                                                                                                                               301
                                                                                                                                                                                                                                                 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      241
```

g

ŏ g ò g ò g ò g ò g ó

à

ö

g

ò g ò δ g δ q ò

```
- malaria parasite (Plasmodium falciparum) (fragments
                                                                                                                                                                                                                                                                                                                                                            protei
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)
N;Alternate names: gp195 surface antigen
C;Species: Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2,
                                                                                                                                                                                                                                                                                                                                                            surface
                                                                                                                                                                                  major merozoite surface antigen - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09.Jun-2000 C;Accession: S05603; S04850 R;Wyler, P.J. submitted to the EMBL Data Library, April 1989 A;Reference number: S05603 A;Accession: S05603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSY 364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                            merozoite
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MDNILSGFENEY DVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   251 NHNNLLTKFLSTGMVFENLAKTVLSNLLDGNLQGMLNISQHQCVKKQCPQNSGCFRHLDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Cross-references: EMBL:X15063; NID:99896; PIDN:CAA33163.1; PID:99897
R;Myler, P.J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 tandem repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 400;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     72; Indels
                                                                                                                                                                                                                                                                R;Blackman, M.J.; Ling, I.T.; Nicholls, S.C.; Holder, A.A. Mol. Blochem. M.J.; Ling, 1.T.; Nicholls, S.C.; Holder, A.A. Mol. Blochem. Parasitol. 49, 29-33, 1991
A;Title: Proteolytic processing of the Plasmodium falciparum A;Reference number: A45545; MUID:92131048
A;Accession: A45545
A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Note: sequence extracted from NCBI backbone (NCBIN:77612, C;Superfamily: major merozoite surface antigen C;Keywords: glycoprotein; merozoite; surface antigen; tande
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 62.5%; Score 1289.5; Dispert Local Similarity 62.3%; Pred. No. 1e-60; Matches 243; Conservative 54; Mismatches 7
                         1668 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 1701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nucleic Acids Res. 17, 5401, 1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: mRNA
A; Residues: 1-1639 <MYL>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Residues: 1-400 <BLA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           365
361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S05603
                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          antigen gene of the malaria parasite Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (s
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1487
                                                                                                                                                              1608 HLDEREECKCLLAYKQEGDKCVENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 1667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEOKKESKFLPFLTNIETL 180
                                                      FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                        HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK
                                                                                                                                 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 1701;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 2006; DB 2; Length Lengt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1668 PDSYPLFDGIFCSSNFLGISFLLIIMLILYSFI 1701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97.3%;
98.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 98.0
Matches 386; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                          1368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                       61
                                                                                                                                 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    301
                                                               g
                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

9. P. οŽ 2

٥y

```
C,Accession: A24594
R;Holder, A.A.; Lockyer, M.J.; Odink, K.G.; Sandhu, J.S.; Riveros-Moreno, V.; Nicholl Nature 317, 270-273, 1985
A;Title: Primary structure of the precursor to the three major surface antigens of Pl A;Reference number: A24594; MUID:86014355
A;Accession: A24594
                                      probable major surface antigen (83K, 19K, 42K) precursor - malaria parasite
C;Species: Plasmodium falciparum
C;Date: 29-Aug-1987 #sequence_revision 29-Aug-1987 #text_change 09-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1550 EREECKCLLNYKOEGSKCVENSNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDC 1609
                                                                                                                  ESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     66 ESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                  |:| : |:|||| |||||| :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVL
                                                                                                                                                                                                                       126 LAKYKDDLESIKKVIKEEKEFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV
                                                                                                                                                                                                                                                                                                                                                                                                                                   246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKF-QDMLNISQHQCVKKQCPENSGCFRHLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1490 NHNNLLTKFLSTGMVFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQOSSGCFRHLD
                                                                                                                                                                                                                                                                                                                              NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKF-QDMLNISQHQCVKKQCPENSGCFRHLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKK ITCECTKPDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 1640;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54.5%; Score 1124; DB 2;
56.3%; Pred. No. 2.2e-51;
ive 60; Mismatches 89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Molecule type: DNA
A;Residues: 1-1640 <40L>
C:Superfamily: major merozoite surface antigen
C;Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1601 YPLFDGIFCSSSNFLGISFLLILMLILYSFI 1631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               YPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                           1381 SEKYKSDLDSIKKYI-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best_Local Similarity
Matches 220; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      246
                                                                                                                                                                                                                                                                                                                              186
                                                                                                                                                                                                                                                                                                                                                                               1421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        304
                                                                                                                  99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                           g
                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ωp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Dp
          ò
                                                                                                                  ò
                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
A; Title: Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene from Plasm A; Reference number: S04850; MUID:89345116
A; Accession: S04850
A; Cossion: S04850
A; Residues: 1504-1639 < MYLL2>
A; Cross-references: EMBL:X15063
A; Cross-references: EMBL:X15063
C; Superfamily: major merozoite; surface antigen
E; 1-19/Domain: signal sequence #status predicted <SIG>E; 20-1639/Product: major merozoite surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SAZORI

SAZORI

SAZORI

SAZORI

SAZORI

SAZORI

SAZORI

S.Species: Plasmodium falciparum

C.Species: Plasmodium falciparum

A.Title: Polymorphism of the precursor for the major surface antigens of Plasmodium falciparum

A.Title: Polymorphism of the precursor for the major surface antigens of Plasmodium falciparum

A.Reference number: A91030; MUID:86136024

A.Residues: 1-1510

A.Residues: 1-1510

C.Comment: Phe merozoite stages of different strains have strain-specific surface antige

C.Comment: Phe merozoite surface antigen

C.Comment: Phe major merozoite surface antigen

C.Superfamily: anajor merozoite; surface antigen

C.Superfamily: anajor merozoite statatus predicted

C.Superfamily: anajor merozoite statatus predicted

C.Superfamily: anajor merozoite statatus predicted

C.Superfamily: anamembrane prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (Asn)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      carbohydrate
                                                                                                                                                                                                                                                                                                                                                                                  5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1390 SEKYKSDLDSIKKYI-------1000GENEKYLPFLNNIETLYKTVN 1429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1550 REECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSY 1609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKY 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    F;20-1631/Product: major merozoite surface antigen #status predicted <MAT>
F;67-84/Region: 3-residue repeats (S-G-T/P)
                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                   6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              305 REECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSY
                                                                                                                                                                                                                                                                                                                                                                                  21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1614-1631/Domain: membrane anchor #status predicted <MBN> 97,259,755,759,835,911,955,1049,1156,1165,1436,1563/Binding site:
                                                                                                                                                                                                                                                                                                                            Length 1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1631;
                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                              DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 85;
                                                                                                                                                                                                                                                                                                                                                                                  72;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             56.5%; Score 1166; DB 1; 58.1%; Pred. No. 1.4e-53;
                                                                                                                                                                                                                                                                                                                         62.5%; Score 1289.5; DB 62.3%; Pred. No. 4.9e-60; iive 54; Mismatches 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 57; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1610 PLEDGIFCSSSNFLGISFLLILMLILYSFI 1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                         Query Match 62.5%
Best Local Similarity 62.3%
Matches 243; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches 227; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
```

365

δ

g

g

126

ò

g

ò

g

Q g δ

99

å

a

g

à

(Plasmodi

ñ

22;

245

ò qq

```
R;Burns Jr., J.M.; Daly, T.M.; Vaidya, A.B.; Long, C.A.
Proc. Natl. Acad. Sci. U.S.A. 85, 602-606, 1988
A;Title: The 3' portion of the gene for a Plasmodium yoelii merozoite surface antigen
A;Reference number: A28121; MUID:88124889
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C;Species: Plasmodium yoelii
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 26-Aug-1999
C;Accession: A28121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        174 LTNIETLYNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHN 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        293 PENSGCFRHLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGK 352
                                                                                                                                                                                                                                                                                                                                                      75 ISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKVLAKYKDDLE 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       309 DILSEFTNESLYVYTKRIGSTYKSLKKHMIREPSTIKEDMTNGLNNKSQKRNDFLEVLSH 368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        369 ELDLFKDLSTNKYVIRNPYQLLDNDKKDKQIVNLKYATKGINEDIETTTDGIKFFNKMVE 428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             KYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKE--SKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            429 VYNTQLAAVKEQIATIE------AETNDTNKEEKKKYIPILEDLKGLYETVI 474
                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   68 DLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKVLA 127
                                                                                                                                                                                                                                              15 NEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVLESDLMQFKH 74
     NCBIP:83592)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                234 DFEAIKKLINDDTKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNISQHQCVKKQC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8 NILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVLES
                                                                                                                                                                                         44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Len
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 680;
                                                                                                                                      Length 1751;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Note: the authors translated the codon GTA for residue 429 as C;Superfamily: major merozoite surface antigen
C;Reywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      major merozoite surface antigen - Plasmodium yoelii (fragment)
                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
A;Note: sequence extracted from NCBI backbone (NCBIN:83591, C;Superfamily: major merozoite surface antigen C;Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Residues: 1-680 <BUR>
A; Cross-references: GB:J03612; NID:g160678; PID:g160679
A; Experimental source: strain 17XL
                                                                                                                                   Query Match 40.3%; Score 832; DB 2; L Best Local Similarity 41.5%; Pred. No. 4.4e-36; Matches 165; Conservative 84; Mismatches 105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      143;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            353 KITCECTKPDSYPLFDGIFCSSSNFLGISFLLILMLIL 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             29.8%; Score 613.5; DB 2 larity 33.5%; Pred. No. 4.2e-25; Conservative 90; Mismatches 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                              135 SIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 131; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Accession: A28121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ολ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                A39401

morozotte surface antigen 1 precursor - Plasmodium vivax

C;Species: Plasmodium vivax

C;Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 24-Nov-1999

C;Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 24-Nov-1999

C;Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 24-Nov-1999

R;del Portillo, H.A.; Longacre, S.; Khouri, E.; David, P.H.

Proc. Natl. Acad. Sci. U.S.A. 88, 4030-4034, 1991

A;Title: Primary structure of the merozoite surface antigen 1 of Plasmodium vivax reveal

A;Reference number: A39401; MUID:91219506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A45604

Maior blood-stage surface antigen Pv200 - Plasmodium vivax
C;Species: Plasmodium vivax
C;Species: Plasmodium vivax
C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
C;Accession: A45604
R;Gibson, H.L; Tucker, J.E.; Kaslow, D.C.; Krettli, A.U.; Collins, W.E.; Kiefer, M.C.; Mol. Blochem. Parasitol. 50, 325-333, 1992
A;Title: Structure and expression of the gene for Pv200, a major blood-stage surface ant A;Reference number: A45604; MUID:92158013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1579 VKSSGLLEKLMKSKLIKENESKEILSQLLNVQTQLLTMSSEHTCIDTNVPDNAACYRYLD 1638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              185 VNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLIND 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       245 DTKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLD 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                75 ISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKVLAKYKDDL- 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     134 ---ESIKKV---IKEEKEKFPSSPPTTPPSPAKTDEQKKES---KFLPFLTNIETLYNNL 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               304 EREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDS 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVLESDLMQFKH 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          40.9%; Score 843; DB 2; Length 17
43.7%; Pred. No. 1.2e-36;
Live 81; Mismatches 115; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1751 <GIB>
A;Cross-references: GB:M75674; NID:g160608; PID:g457336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Molecule type: DNA
A; Residues: 1-1726 < DEL>
A; Cross-references: GB: M60807
C; Superfamily: major merozoite surface antigen
C; Keywords: surface antigen
                                                        1610 YPLSMVIFCSSSNFLGISFLLILMLILYSFI 1640
                              YPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1695 EPLFEGVFCSSSFLSLSFLLMLFL 1721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          364 YPLFDGIFCSSSNFLGISFLLILMLIL 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ouery Match
Best Local Similarity 43.73
Matches 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Accession: A45604
                            364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     15
```

ò g à a g

ò

合

ά g οy

δ

g

g

ò

ί,

ø

```
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum. A;Reference number: A71600; MUID:99021743
A;Accession: E71606
A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA
A;Residues: 1-980 <GAR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1532 MIELYKPOLNAVNEQI-----AAIGTEP----TDAEKK--KYAPIFEDLKGLYETI 1576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1630 HIASIALNNLNKSGLVTEGESKKILAKMLNMDAMDLLGIGSNHVCISTSTPDNAGCFRYD 1689
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C.Species: Plasmodium falciparum
C.Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               65 LESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEK 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   125 VLAKYKDDLESIKKVIKEEKERFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNL 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEREECKCLLNYKQ - - EGDKCVENPNPTCNENNGGCDADAKCTEEDSGSN - - GKKITCEC 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8 NILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDI---LNSRLKKRKYFLDV 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C;Accession: E71606
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, Science 282, 1126-1132, 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hypothetical protein PFB0765w – malaria parasite (Plasmodium falciparum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     185 VNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DTKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNI-SQHQCVKKQCPENSGCFRHL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28.3%; Score 583; DB 2; Lv 33.2%; Pred. No. 4.8e-23; Live 92; Mismatches 137;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      359 TKPDSYPLFDGIFCSSSNFLGISFLLILMLILYS 392
                       1740 TPNAYYEGVFCSSSSFMGLSILLITLIVFN 1770
  362 DSYPLFDGIFCSSSNFLGISFLLILMLILYS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 28.3
Best Local Similarity 33.2
Matches 131; Conservative
                                                                                                         14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1690
                                                                                                           RESULT
                                         g
  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   á
                                                                                                                                                                                                                                                                                                                                                                    A45532
major merozoite surface antigen precursor - Plasmodium yoelii
(;Species: Plasmodium yoelii
(;Species: Plasmodium yoelii
(;Accession: A45532; A4531
R;Lewis, A.P.
Rol. Biochem. Parasitol. 36, 271-282, 1989
A;Title: Cloning and analysis of the gene encoding the 230-kilodalton merozoite surface A;Accession: A45532
A;Accession: A45532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: DNA
A; Residues: 1-1772 < LEWN>
A; Residues: 1-1772 < LEWN>
A; Cross-references: GB: J04668; NID:g160492; PID:g160493
R; Daly, T.M.: Burns Jr., J.M.; Long, C.A.
Mol. Biochem. Parasitol. 36, 283-285, 1989
A; Title: Precursor to the major merozoite surface antigen of Plasmodium yoelii: cloning
A; Reference number: A45531; MUID:90014982
A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1401 DILSEFTNESLYVYTKRLGSTYKSLKKHMLREFSTIKEDMTNGLNNKSQKRNDFLEVLSH 1460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1620 IASIALNNLNKSGLVGEGESKKILAKMINMDGMDLLGVDPKHVCVDTRDIPKNAGCFRDD 1679
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           647
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKVLA 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        186 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNIS-QHQCV-KKQCPENSGCFRHL 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEREECKCLLNYKQ-EGDKCVENPNPTCNENNGGCDADARCTEEDSGSNGKKITCECTKP 361
                                                                                                                           528 IASIALNNLNKSGLVGEGESKKILAKMLNMDGMDLLGVDPKHVCVDTRDIPKNAGCFRDD 587
                                                                                                                                                                    303 DEREECKCLLNYKQ-EGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKP 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVLES 67
                                                                                                                                                                                           186 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD
                                      475 GQAEEYSEELQNRLDNYKNEKAEFEILTKNLEKYIQIDEKLDEFVEHAE-----NNKH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    128 KYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKE--SKFLPFLTNIETLYNNLV
                                                                                  246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNIS-QHQCV-KKQCPENSGCFRHL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1772;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule, type: DNA
A; Residues: 454-1094 <DAL>
A; Cross-references: GB:J03975; NID:g160081; PID:g160082
C; Superfamily: major merozoite surface antigen
C; Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        . 2e-24;
nes 143;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29.8%; Score 613.5;
33.5%; Pred. No. 1.2e
Live 90; Mismatches
                                                                                                                                                                                                                                                                         648 TPNAYYEGVFCSSSSFMGLSILLIITLIVFN 678
                                                                                                                                                                                                                                                    362 DSYPLFDGIFCSSSNFLGISFLLILMLILYS 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best_Local Similarity 33.5
Matches 131; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      89
  ð
                                      a
                                                                                                                    g
                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                  à
                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

÷

11;

34;

Length 1785; Indels

```
A;Cross-references: GB:AE001417; GB:AE001362; NID:g3845271; PIDN:AAC71942.1; PID:g384527
A;Experimental source: clone 3D7
Genetics (Genetics)
A;Gene: PPB0765w
C;Superfamily: hypothetical protein MJ1322
                                                                                                                                                                             23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   139 -----VIKEEKEKFPSSPPTTPPSPAKTDEQ-----KKESKFLPFLTNIETLYN-NLVNKI 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  : :|:| |::|::| |:: | 302 KDRKIDLLTNIENELLKKKEINNIKLMEKQNVIKNNEQLLKDIKDENEKMNEHVNKLQNE 361
                                                                                                                                                                                                                                                                                                                                                                        LLNSEQKNTLL-KSYKYIKESVENDIKFAQEGISYYEKVLAKYKDDLESIKK----- 138
                                                                                                                                                                                                                                                                                                                                                                                                |: | : | : | : |: |: || 362 LIKRELQNKCISKDIEFCKKEKEBKIKNLEDDLLEKKKCIENLKDELINIKKKMEDKMHM 421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           422 TNEMDLLSNKVEELNRINKTYEKNIVELNNELDVIKKKLNDEEFLKEEFKKKINDWYKI 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             189 DDYLINLKAKINDC-NVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDDTK 247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     248 -----KDMLGKLLSTGLVQNFPNTIISKLIEGKFQ------DMLNISQHQCVKKQCPE 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      295 NSGCFRHLDER-EECKCLLNYK-----QEGDKCVENPPTCNENNGGCDADAKCTEE 345
                                                                                                                                                                                                                                         14 ENEYDVIYLKPLAGVYRSLKKQIEKNI-----FTFNLNLNDIL---NSRLKK 57
                                                                                                                                  Query Match 8.8%; Score 180.5; DB 2; Length 980; Best Local Similarity 24.0%; Pred. No. 0.028; Matches 103; Conservative 75; Mismatches 141; Indels 111;
                                                                                                                                                                                                                                                                                            58 RKYFLDVL---ESDLMQFKHIS-----SNEYIIED-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 346 DSGSNGKKIT 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         : | ||:|
645 QAKINNKKLT 654
                                                                                                                                                                                                                                                                                                                                                                        88
                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                            oy
Op
                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

Search completed: August 8, 2001, 12:30:49 Job time: 97 sec

THIS PAGE BLANK (USPTD)

```
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

August 8, 2001, 12:29:12; Search time 24.79 Seconds (without alignments) 544.440 Million cell updates/sec Run on:

US-09-500-376-2 2062 1 AISVTMDNILSGFENEYDVI......SNFLGISFLLILMLILYSFI 394 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

93435 segs, 34255486 residues Searched:

93435 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_39:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	•			P08569 plasmodium				P13828 plasmodium						Q13201 homo sapien		_	P54697 dictyosteli	P43573 saccharomyc		Q00799 plasmodium		P17119 saccharomyc	P41004 schizosacch	P12753 saccharomyc	P39723 saccharomyc	P47867 mus musculu	Q02455 saccharomyc		P51819 pharbitis n	9		190	P38989 saccharomyc
SUMMARIES	OI	MSP1_PLAFC	MSP1_PLAFP	MSP1_PLAFF	MSP1_PLAFM	MSP1_PLAF3	MSP1_PLAFK	MSP1_PLAFW	MSP1_PLAYO	YL17_CAEEL	SMC1_YEAST	MYS2_DICDI	RBP1_PLAVB	YIO9_YEAST	ECM_HUMAN	USO1_YEAST	NUF1_YEAST	MYSJ_DICDI	YFC3_YEAST	CIN8_YEAST	RBP2_PLAVB	YKZ6_YEAST	KAR3_YEAST	CUT3_SCHPO	RA50_YEAST	YAE7_YEAST	SG3_MOUSE	MLP1_YEAST	- 1	HS83_PHANI	ITB1_HUMAN	- 1	٩,	SMC2_YEAST
	DB	-	~	٦	П	-	1	-	٦	-	-	<del>, -1</del>	Н	-	-1	-	-	Н	-	<b>~</b>	-	Н	Н	-	-		Н	-		<del>-</del>	-	-		-
	Query Match Length	,	1726	1701	1701	1682	1630	1639	1772	1130	1225	2116	2869	1679	1228	1790	944	2245	196	1038	1251	540	729	1324	1312	622	471	1875	1928	703	798	1230	399	1170
de	Query Match	100.0	99.9			98.2					•		•	9.9	•	•	•		•	•	6.2		•	•					٠		٠		5.9	٠
	Score		2059	2027	2027	2025.5	1289.5	1289.5	613.5	153	139.5	137	136.5	. 136	135	130.5	129	129	128	128	128	127	127	127	12		125	٠	CV.				122	C)
	Result		7	m	4	5	9	7	80	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	53	30	31	32	33

P47460 mycoplasma P06101 saccharomyc	P16320 drosophila P25805 plasmodium	P07228 gallus gall	Q9usi6 schizosacch	Q99104 mus musculu	P22620 plasmodium	Q01969 thermotoga	P23745 plasmodium	P38822 saccharomyc
HMW2_MYCGE CC37_YEAST	NOF_DROME CYSP_PLAFA	ITB1_CHICK	MYS2_SCHPO	MY5A_MOUSE	ABRA_PLAFC	OMPA_THEMA	ABRA_PLAFG	YHR4_YEAST
		٦-	-	-1	~	7	٦	7
1805 506	984 569	803 1101	1526	1853	743	400	009	633
0.0 0.0	0.50 0.00	ი ი ი ი	8.0	œ.	5.8	5.8	5.8	5.8
122 121.5	121.5 121	121	120.5	120.5	120	119.5	119.5	119.5
34 35	36 37	8 6 8 8	40	41	42	43	44	45

## ALIGNMENTS

ANNUAL CONTRACT TARKED AND CONTRACT TO CONTRACT TO CONTRACT TARKED AND CONTRACT TARKED TARKED TARKED TARKE
--

```
RESULT 3
MSP1_PLAFF
1D MSP1_PLAFF
AC P13819;
DT 01-JAN-199'
DT 01-JAN-10
DT 01-OCT-
DE MEROZC
DE (PMM'
GN MSP
OS P'
                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   301
      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               s SWISS-PROT entry is copyright. It is produced through a collaboration ween the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                       1333 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRY 1392
                                                                                                                                                                                                                                                                                                                                                                        1452
                                                                                                                                                                                                                                                                                                                                                                                                                                  1632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          There are no restrictions on
                                                                                                                                                                                                                                                                                                                                                                                                            180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             360
                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          isolate.";
Exp. Parasitol. 67:1-11(1988).
-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
01-0CT-1996 (Rel. 34, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                   YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HLDEREECKCLLNYK@EGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LINKED (GLCNAC. .) (POTENTIAL).
DD8AD45FA352BCF3 CRC64;
                                                                                                                                                                                                                                                       AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                   FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hydropathy profile of
the Uganda-Palo Alto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL).
PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                 ö
                                                                                                                                                                          1; Length 1726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   / Uganda).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-89005525; Pubmed-3049134;
Chang S.P., Kramer K.J., Yamaga K.M., Kato A., Case S.E.,
                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (isolate Palo Alto / Uganda
Apicomplexa; Haemosporida;
                                                                                                                                                                                                                 ö
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
                                                                                                                                                                        100.0%; Score 2062; DB 1
100.0%; Pred. No. 9e-99;
iive 0; Mismatches 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PDSYPLEDGIFCSSNFLGISFLLILMLILYSFI 1726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "Plasmodium falciparum: gene structure and major merozoite surface antigen (gp195) of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1726
                                                                                                                    ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "Plasmodium falciparum: gene
                                                                                                                  AA; 196197
                                                                                                                                                                                                                 Conservative
  944
990
1016
11114
1221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Plasmodium falciparum
Eukaryota; Alveolata;
NCBI_TaxID=57270;
                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (PMMSA) (GP195)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Siddigui W.A.;
                                                                                                                                                                                                                 394:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MSP1_PLAFP
P50495;
                                    CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
SEQUENCE
CARBOHYD
CARBOHYD
                                                                                                                                                                          Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 2
MSP1_PLAFP
AC P50495;
DT 01-0CT-
DT
                                                                                                                                                                                             Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                      1393
                                                                                                                                                                                                                                                                                                                                                                                                                                                1453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1693
                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           301
                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
  FTFFF
                                                                                                                                                                                                                                                                                                                               δy
                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
```

```
entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1333 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY 1392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1453 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPARTDEQKKESKFLPFLTNIETL 1512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1633 HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 1692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1393 FLDVLESDLMQFKHISSNRYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (Rel. 13, Created)
(Rel. 13, Last sequence update)
(Rel. 34, Last anotation update)
JRFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFAÇE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .;
0
                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                     Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 2059; DB 1; Length 1726;
Pred. No. 1.3e-98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Plasmodium falciparum (isolate FC27 / Papua New Guinea).
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
        content
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                    PROTEIN 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5B59CEEFA2F9A026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
                                                                                                                                                                                                                                                                                              (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                  SURFACE
                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                       (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1693 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 1726
use by non-profit institutions as long a modified and this statement is not removed. entities requires a license agreement (See h
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; Mismatches
                                                                                                                                                                                                                  MEROZOITE S
N-LINKED (C
                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              196174
                                                                                                EMBL; M37213; AAA29611.1;
InterPro; IPR000561; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches 393; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                     1114
1221
1613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JAN-1990 (Rel. 1
01-JAN-1990 (Rel. 1
01-OCT-1996 (Rel. 3
MEROZOITE SURFACE P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
```

```
Tanabe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIGNAL
                            361
                                                                                                            MSP1_PLAFM
                                                       1668
                                                    용
                         ò
                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FLDVLESDLAQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 1427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PTM: MEROZÓITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                     MEDLINE-88142999; PubMed=2449612;
Peterson M.G., Coppel R.L., McIntyre P., Langford C.J., Woodrow G., Brown G.V., Anders R.F., Kemp D.J.;
"Variation in the precursor to the major merozoite surface antigens of Plasmodium falciparum.";
                                                                                                     Biochem. Parasitol. 27:291-302(1988).
SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                          InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 2027; DB 1; Length 1701;
Pred. No. 5.5e-97;
1; Mismatches 4; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NKED (GLCNAC. . .) (POT 3920B75E73D38552 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                              SURFACE
                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC. N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                              MEROZOITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MM;
                                                                                                                                                                                                                                                                                                              EMBL; M19143; AAA29653.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        193719
                                                                                                                                                                                                                                                                                                                                                                                   GPI-anchor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ¥
                                                                                                                                                                                                                                                                                                                              PIR; A54498; A54498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
NCBI_TaxID=5837;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1089
1196
1588
1701
                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                   Pransmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          389;
                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local
Matches 38
                                                                                                                                                                                                                                                                                                                                                                                              SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  301
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
1608 HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 1667
                                                                                                                                                                                                                                                                                                                                                                               parasite
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL).
PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 4 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1-115 FROM N.A. MEDLINE-86136024; PubMed-3004972; MacKay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U., Stunnenberg H., Bujard H.; Pujard H.; Polymorphism of the precursor for the major surface antigens of plasmodium falciparum merozoites: studies at the genetic level."; EMBO J. 4:3823-3829(1985).
                                                                                                                                                                       30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
(PMMSA) (P190)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                          MEDLINE-88011243; PubMed-3079521;
Tanabe K., Mackay M., Goman M., Scaife J.G.;
"Allelic dimorphism in a surface antigen gene of the malaria
                                                                                                                                                                                                                                                                     Guinea).
Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PROTEIN 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                   (isolate mad20 / Papua New Apicomplexa; Haemosporida;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.

MEROZOITE SURFACE

N-LINKED (GLCNAC.

N-LINKED (GLCNAC.
                                                 PRT; 1701 AA
                                PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                          [2]
REVISIONS TO 1403; 1569 AND 1629.
                                                                                                                                                                                                                                                                                                                                                                                                             Mol. Biol. 195:273-287(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; X05624; CAA29112.1; -.
                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                   Plasmodium falciparum
Eukaryota; Alveolata;
                                                                                                                                                                                                                                                                                                                                                                                             Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR000561;
Pfam; PF00008; EGF; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PIR; A26868; A26868
PIR; B25120; B25120
                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                  NCBI_TaxID=70153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Fransmembrane;
                                                                                                                                        MSP1_PLAFM
P08569;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
```

```
PIR; S06286; S06286.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-AUG-1987 (Rel. 0
01-FEB-1996 (Rel. 3
01-OCT-1996 (Rel. 3
MEROZOITE SURFACE P
                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (PMMSA) (P190).
                                                                                                                                                            Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 6
MSP1_PLAFK
ID MSP1_PLAFK
AC P04932;
                                                                                                                                                                                              TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local
                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                     SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             301
   Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PTM: MEROZÓITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42
KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                         1308 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY 1367
                                                                                                                                                                     FLDVLESDLAQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 1427
                                                                                                                                                                                                                   240
                                                                                                                                                                                                                                                                                                     300
                                                                                                                                                          FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                        YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                   HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 360
                                                                                    Gaps
                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-80166657; PubMed=3327688;
Certa U., Rotmann D., Matile H., Reber-Liske R.;
"A naturally occurring gene encoding the major surface antigen
precursor p190 of Plasmodium falciparum lacks tripeptide repeats.";
EMBO J. 6:4137-4142(1987)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EXP. Parasitol. 81:47-54(1995).
-i- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                           1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                    241 LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                       YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ō
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Tolle R., Bujard H., Cooper J.A.; "Plasmodium falciparum: variations within the C-terminal region
            N-LINKED (GLCNAC. . .) (POTENTIAL)
MW; 3FC2EC59AF96EA98 CRC64;
                                                                                    ő
                                                           Score 2027; DB 1; Length 1701;
Pred. No. 5.5e-97;
1; Mismatches 4; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Plasmodium falciparum (isolate ro-33 / Ghana).
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
NCBL_TaxID-5834;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SURFACE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MSP1_PLARF3 STANDARD; PRT; 1682 AA.
P19598; 025921;
01-FEB-1991 (Rel. 17, Created)
01-NVV-1997 (Rel. 35, Last sequence update)
01-NVV-1997 (Rel. 35, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE;
  N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                             PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 1032-1682 FROM N.A. MEDLINE-95354793; PubMed-7628566;
1196 1196
1588 1588
1701 AA; 193768 M
                                                            98.3%;
98.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 1-1061 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         merozoite surface antigen-1
                                                          Query Match
Best Local Similarity 98.7
Matches 389; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (PMMSA) (P190).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
 CARBOHYD
              CARBOHYD
                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 5
MSP1_PLAF3
                                                                                                                                                                                 1368
                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                           1548
                                                                                                                                                                                                                                                                                                                                                  301
                                                                                                                                                                                                                                                                                                                                                                         1608
                                                                                                                                                                                                                                                                                                                                                                                                 361
                                                                                                                                                         61
                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                                        1668
                                                                                                                                                                                                                                                                                                                                                                                                                    8
  F F 8
                                                                                                                                q
                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                         à
                                                                                                                                                                          g
                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                             ò
```

```
ons on its
in no way
commercial
                                   use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1290 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY 1349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       241 LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
   restrictions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         05, Created)
33, Last sequence update)
10. Last annotation update)
PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              D (GLCNAC. . . ) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Length 1682;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
There are no restrict of as its content is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL.
MEMBRANE ANCHOR (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTEN N-LINKED (GLCNAC. . .) (PO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 2025.5; DB 1; Length
Pred. No. 6.5e-97;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CRC64:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C82A1E159948CAD6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1649 PDSYPLEDGIFCSSSNFLGISFLLILMLILYSFI 1682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1630 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ψ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19
20 1682
566 1682
33 233
62 858
9 599
5 785
901
1071
1178
1569
                                                                                                                                                                                                                              EMBL; M35727; AAA29715.1; -. EMBL; Y00087; CAA68280.1; -. EMBL; Z35326; CAA84555.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98.2%;
99.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     390; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
```

ij

```
InterPro; IPR000561;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIR; A24594; A24594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                         (PMMSA) (P195).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Fransmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Freeman R.R.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Holder A.A.;
                                                                                                                                                                                                                                                                             13-AUG-1987
                                                                                                                                                                                                                                                  MSP1_PLAFW
P04933;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REVISIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIGNAL
                                                                                                                                                                                                                                    MSP1_PLAFW
                              ò
                                                        g
                                                                                                                 g
                                                                                                                                                ð
                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1381 SEKYKSDLDSIKKYI--------NDKQGENEKYLPFLNNIETLYKTVN 1420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          186 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                 Q.
                                                                                                                                                                                      Pan W., Tolle R., Bujard H.;
Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                             PTM: MEROZÓITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                     SEQUENCE FROM N.A.
MACKAY M., Gomen M., Bone N., Hyde J.E., Scaife J., Certa U.,
Mackay M., Gomen M., Bone N., Hyde J.E., Scaife J., Certa U.,
Stunnenberg H., Bujard H.;
Stunnenberg H., Bujard H.;
Polymorphism of the precursor for the major surface antigens of
Plasmodium falciparum merozoites: studies at the genetic level.";
EMBO J. 4:3823-3829(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 62.5%; Score 1289.5; DB 1; Length 1630;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBL_TaxID=5839;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              72; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADBDEC3CE0A46322 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEROZOITE SURFACE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. 3.5e-59;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRIPEPTIDE SG(TP)
.asmodium falciparum (isolate K1 / Thailand)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEMBRANE ANCHOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              54; Mismatches
                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; X03371; CAA27070.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       187289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GPI-anchor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity was Matches 243; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1630
84
1630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1156
1165
1436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            049
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; A25120; SAZQK1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1165
1436
                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ransmembrane;
                                                                                                                                                            [2]
REVISIONS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99
```

oy Op ò

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V., Nicholls S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PTM: MEROZÓITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, «KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Submitted (MAR-1991) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               13.AUG-1987 (Rel. 05, Created)
1-FEB-1996 (Rel. 33, Last sequence update)
01.FEB-1996 (Rel. 33, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                     246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDE
                                                                                                                                                                           305 REECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Primary structure of the precursor to the three major surface antigens of Plasmodium falciparum merozoites.";
Nature 317:270-273(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Plasmodium falciparum (isolate Wellcome).
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBL_TaxID=5848;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PROTEIN 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEROZOITE SURFACE
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                               1639 AA.
                                                                                                                                                                                                                                                                                                           1601 PLFDGIFCSSSNFLGISFLLILMLILYSFI 1630
                                                                                                                                                                                                                                                                                 365 PLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
MEDLINE-86014355; PubMed-2995820;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; X02919; CAA26676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GPI-anchor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
```

ø

us-09-500-376-2.rsp

```
YL17_CAEEL
Q11102;
01-NOV-1997 (
01-NOV-1997 (
15-JUL-1998 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 9
YL17_CAEEL
ID YL17_C3
AC Q11102
AC Q11102
DT 01-NOV
DT 15-JUL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             128
    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1489
                                                                                                                                                                                                                                         1550 REECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSY 1609
                                                                                                                                                                                                                                                                                                                                                                                                                                                           245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  364
                                                                                                                                                                                                                                                                                                 ESDLMOFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Sequence analysis upstream of the gene encoding the precursor to the major merozoite surface antigens of Plasmodium yoelii."; Mol. Biochem. Parasitol. 39:285-288(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDE 304
                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                       MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 4 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Proc. Natl. Acad. Sci. U.S.A. 85:602-606(1988).
-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                           186 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-JAN-1990 (Rel. 13, Created)
01-JAN-1990 (Rel. 13, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                               LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDSY
                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Burns J.M. Jr., Daly T.M., Vaidya A.B., Long C.A.; "The 3' portion of the gene for a Plasmodium yoelii merozoite sur antigen encodes the epitope recognized by a protective monoclonal
                                                                                                                                           DB 1; Length 1639;
                                                                                                                                                                                21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                Indels
                                      N-LINKED (GLCNAC. . .) (POT
N-LINKED (GLCNAC. . .) (POT
W; 2C255B6616C87F6E CRC64;
                                                                                                                                       Query Match 62.5%; Score 1289.5; DB 1; Best Local Similarity 62.3%; Pred. No. 3.5e-59; Matches 243; Conservative 54; Mismatches 72;
  (GLCNAC...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT; 1772 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1610 PLFDGIFCSSSNFLGISFLLILMLILYSFI 1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PLFDGIFCSSSNFLGISFLLILMLILYSFI 394
    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A. MEDLINE-90205979; Pubmed-2320061; Lewis A.P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-88124889; PubMed-2448778;
                                                                               ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 1093-1772 FROM N.A.
                                                                               187618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           yoelii
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
  1165
1174
1445
1526
                                                                                  AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Plasmodium berghei
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (PMMSA) (230 KDA).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL).
  1165
1174
1445
1526
1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=5862;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-17XL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MSP1_PLAYO P13828;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota;
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             antibody.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MSPI_PLAYO

TO MSPI_PLAYO

TO 01-JAN

DT 01-JAN

DT 01-JAN

DT 01-JAN

DE MEROZO

DE MEROZO

OC EUKATY

OC EUK
                                                                                                                                                                                                                                                                                                                                                                               126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              365
                                                                                                                                                                                                                         ٥
                                                                                                                                                                                                                                                                                                   99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                305
    FFFS
                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                     modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      :| || :|:|:|:| : |:|:|:|: || : : || : : || : : ||: : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1620 IASIALNNLNKSGLVGEGESKKILAKMLNMDGMDLLGVDPKHVCVDTRDIPKNAGCFRDD 1679
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDD 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            303 DEREECKCLLNYKQ-EGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKP 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKVLA 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              27; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8 NILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDVLES 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               246 TKKDMLGKLLSTGLV-QNFPNTIISKLIEGKFQDMLNIS-QHQCV-KKQCPENSGCFRHL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKE--SKFLPFLTNIETLYNNLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 1772;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
MEROZOITES, THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  . V (IN REF. 2).
9A6291658EB0F45D CRC64.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          29.8%; Score 613.5; DB 1; 33.5%; Pred. No. 1.8e-24; ive 90; Mismatches 143;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.
MENOZOITE SURFACE PI
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1740 TPNAYYEGVFCSSSFMGLSILLITLIVFN 1770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DSYPLFDGIFCSSSNFLGISFLLILMLILYS 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (Rel. 35, Created)
(Rel. 35, Last sequ
(Rel. 36, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; J03612; AAA29762.1; -.
EMBL; J04668; AAA29702.1; -.
PIR; A28121; A28121.
PIR; A45532; A45532.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1521
AA; 197230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GPI-anchor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 33.5
Matches 131; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         646
829
1018
1090
1408
1541
1529
1680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1090
1408
1446
1541
1629
1680
1721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Transmembrane;
SIGNAL
```

```
MEDLINE=94103320; PubMed=8276886;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NP_BIND
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MUTAGEN
MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          378
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           863
                                                                                                                                                                                                                                                                                                                                                                                                                                                    81 IIEDSFKLLNSEQKN-TLLKSYKYIKESVENDIKFAQEGISYYEKVLAKYKDDLESIKKV 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  140 IKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLVNKIDDYLINLKAKI 199
                                                                                                                                                                                                                                                                                                                                                                                                                           596 NLRKOLEKEI-SHTEDRNRLLHENTOKELEAHKETHTETVRVLEAEIDQFKSAFENEQEY 654
           Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
Rhabditidae; Peloderinae; Caenorhabditis.
NCBI_TaxID=6239;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SLKKQIEKNIFTFNLNLNDILNSRLKK------RKYFLDVLESDLMQFKHISSN--EY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  258 GLVQNFPNTIISKLIEGKFQDMLNISQ-----HQCVKKQCPENSGCFRHLDE-----REE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 200 NDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKL--INDDTKKDMLGKLLST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          762 DQMGTIKDEL---VKKNEEIKTISAKTAQLLESNTVESETKLASVTEEREKE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Fung1; Ascomycota; Saccharomycotina; Saccharomycetes;
                                                                                                                                                                                                                                                                                                                                                                               90;
                                                                                                                                                                                                                                                                                                                                                     Length 1130;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-071-1993 (Rel. 27, Created)
01-071-1993 (Rel. 27, Last sequence update)
15-JUL-1998 (Rel. 35, Last annotation update)
CHROWOSME SEGREGATION PROTEIN SMC1 (DA-BOX PROTEIN SMC1).
SMC1 OR CHLJO OR YFLOO8W.
                                                                                                                                                                                                                                                    EMBL; U41545; AAA83190.1; -.
WormPep; C02F12.7; CE03901.
Hypothetical protein; Coiled coil.
DOMAIN 1279 COILED COIL (POTENTIAL).
DOMAIN 805 1061 COILED COIL (POTENTIAL).
SEQUENCE 1130 AA; 131485 MW; BOFDZEFE3D99FB09 CRC64;
                                                                                                                                                                                                                                                                                                                                                    7.4%; Score 153; DB 1; Length 11
24.1%; Pred. No. 0.46;
tive 64; Mismatches 125; Indels
                                                                                                                    Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases -! - SIMILARITY: WEAK, TO MYOSINS.
HYPOTHETICAL 131.5 KDA PROTEIN C02F12.7 IN CHROMOSOME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBL_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1225 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Saccharomyces cerevisiae (Baker's yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 308 CKCLL--NYKQEGDKCVENPNPTCNENN 333
                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
les 79; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                               SEQUENCE FROM N.A.
                                                                                             STRAIN-BRISTOL N2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SMC1_YEAST
P32908;
                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local S:
Matches 79
                                                                                                         Miller N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 10
SMC1_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                     31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
Strunnikov A.V., Larionov V.L., Koshland D.; "SMC1: an essential yeast gene encoding a putative head-rod-tail protein is required for nuclear division and defines a new ubiquitous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATP (POTENTIAL).
COILED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
ALA/ASP-RICH (DA-BOX).
S->L: IN TS MUTANT SMC1-2.
N->D: IN TS MUTANT SMC1-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         86 FKLLNSEQKNTLLKSYKYIKESVENDÍKFAQEGISYYEKVLAKYKDDLESIKKVIKEEKE 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 146 KFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLVNKIDDYLINLKAKINDCNVE 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --FEAIKKLIND--DTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCV-- 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 127; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           329 IEKRI-----ESLQKDLQRQKTYVERFET---QLKVVTRSKEAFEEEIKQSARNYDK 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             36 IEKNIFTFNLNLNDILNSRLKKRKYFLDVLESDLMQFKHISSNEYIIE------DS 85
                                                                                                                                                                                                                                                                                                                                                                                                                                       Nat. Genet. 10:261-268(1995).

-!- FUNCTION: INVOLVED IN CHROMOSOME SEGREGATION IN MITOSIS. COULD IN PART OF A CHROMOSOME CONDENSATION MOTOR.

-!- SUBJUIT: HOMODIMER OF HETERODIMER WITH SMC2 OR OLIGOMERS.

-!- SUBJUIT: HOMODIMER OF HETERODIMER WITH SMC2 OR OLIGOMERS.

-!- SUBJUIT: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS FLANKED BY PUTATIVE GLOBULAR REGIONS AT THE N- AND C-TERMINUS.

-!- SIMILARITY: BELONGS TO THE SMC FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       464 HTERLHELKKLQSDIESANNQEYDLNFKLRETLVKIDDLSANQRETMKERKLRENIAMLK
                                                                                                                                                                                                                                                                                                                 Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
Yamazaki M., Tashiro H., Eki T.;
"Analysis of the nucleotide sequence of chromosome VI from
                                                                                                                                                                                                                                                                                          Ozawa M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       B504017AA0ECCA8C CRC64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                206 KDEAHVKITKL-SDLKAIDD-----KI-DLFKNHND----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    coil; Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6.8%; Score 139.5; DB 1; 22.9%; Pred. No. 2.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 62; Mismatches 110;
                                                                                                                                                                                                                                                                                   Murakami Y., Naitou M., Hagiwara H., Shibata T.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             or send an email to license@isb-sib.ch).
                                                                                               protein family.";
J. Cell Biol. 123:1635-1648(1993).
                                                                                                                                                                                                                  STRAIN=S288C / AB972;
MEDLINE=95400292; PubMed=7670463;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; L00602; AAA16595.1; -. EMBL; D50617; BAA09230.1; -. EMBL; D31600; BAA06496.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Coiled
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          141279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                Saccharomyces cerevisiae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              489
489
1063
1061
1164
173
458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PIR; A49464; A49464.
PIR; S41804; S41804.
SGD; S0001886; SMC1.
Mitosis; ATP-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
Matches 89; Conserv
                                                                                                                                                                                       SEQUENCE FROM N.A.
```

21;

œ

Biochemistry 36:11619-11628(1997)

```
MEDLINE-95345066; PubMed=7619795;
Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
                                      "Replacement of threonine residues by serine and alanine in a
phosphorylatable heavy chain fragment of Dictyostelium myosin II.";
FEBS Lett. 269:239-243(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Dictyostelium discoideum myosin motor domain to 1.9-A resolution."; Biochemistry 35:5404-5417(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-97452580; PubMed-9305951; Gullick A.M., Bauer C.B., Thoden J.B., Rayment I.; Gullick A.M., Bauer C.B., Thoden J.B., Rayment I.; X-ray structures of the MgADP, MgATPgammaS, and MgAMPPNP complexes of the Dictyostellum discoideum myosin motor domain.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Smith C.A., Rayment I.; "X-ray structure of the magnesium(II)-pyrophosphate complex of the truncated head of Dictyostelium discoideum myosin to 2.7-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        'X-ray structure of the magnesium(II).ADP.vanadate complex of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-88112226; PubMed-2828113;
Wagle G., Noegel A., Scheel J., Gerisch G.;
Phosphorylation of threonine residues on cloned fragments of
Dictyostelium myosin heavy chain.";
FEBS Lett. 227:71-75(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "X-ray structures of the myosin motor domain of Dictyostelium discoideum complexed with MgADP.BeFx and MgADP.AlF4-."; Biochemistry 34:8960-8972(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-87092266; PubMed-3540939; Warrick H.M., de Lozanne A., Leinwand L.A., Spudich J.A.; "Conserved protein domains in a myosin heavy chain gene from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Lueck-Vielmeter D., Schleicher M., Grabatin B., Wippler J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Dictyostelium discoideum (Slime mold).
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
MEDLINE-95345067; Pubmed=7619796;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
MEDLINE-96206189; PubMed-8611530;
                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1988 (Rel. 09, Created)
01-OCT-1989 (Rel. 12, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
MYOSIN II HEAVY CHAIN, NON MUSCLE.
                                                                                                                                                                                                                                                                                                                                                                   2116 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PHOSPHORYLATION SITES, AND MUTAGENESIS.
                                                                                                                               330 NENNGGCDADAKCTEEDSGSNGKKITCE 357
                                                                                                                                                                                                 NTLN----IAKDLKWKKGIRGKLVTIE 655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-AX2;
MEDLINE-90353583; PubMed-2387408;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Biochemistry 34:8973-8981(1995).
-KKOCPENSGCFRHLDERE-ECKCL-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Dictyostellum discoideum.
                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Smith C.A., Rayment I.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PHOSPHORYLATION SITES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=44689;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      resolution."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Rayment I.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gerisch G.;
                                                                                                                                                                                                                                                                                                                                                                MYS2_DICDI
                                                                                                                                                                                                                                                                                                                                                                                                   P08799;
                                                                                                                                                                                                                                                                                                  RESULT 11

MYS2_DICDI

DD MYS2_D

DC DASS2_DICDI

DT 01-079

DT 01-079

DT 01-071

DT 01
                                                                                                                                                                                                 633
                                                       g
                                                                                                                               ò
                                                                                                                                                                                              셤
```

```
CORTEX.

MEROMYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (HMM). IT CAN BE FURTHER SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (SI) AND 1 ROD-SHAPED

SUBFRAGMENT (S2).

CORTES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CARRACTERISTIC FOR ALPHA-HELICAL COILED COILS.

CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

THE ACTIN-ACTIVATED ATPASE ACTIVITY.

MISCELLANBOOMS. DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1 POSITION (689).

CIPTIN PAINS BY COMPREATED WITH THE ABSENCE S NONMUSCLE SIMILARITY HITH THE GLOBULAR HEAD SEQUENCE SOWNERVED, BUTTHE PERIODICITIES OF HYDROPHOSIC & CHARGED RESIDUES, WHICH DICTATE THE PERIODICITIES OF HYDROPHOSIC & CHARGED RESIDUES, WHICH DICTATE THE PERIODICITIES OF HYDROPHOSIC & CHARGED RESIDUES, WHICH DICTATE THE PERIODICITIES OF HYDROPHOSIC & CHARGED RESIDUES, WHICH DICTATE THE ALPHA-HELICAL COILED-COIL STRUCTURE, ARE CONSERVED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions a so its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                            (MLC-2).
SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
                      X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
MEDLINE=98070605; PubMed=9405148;
Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
X-ray crystal structure and solution fluorescence characterization of Mg.2'(3')-0-(N-methylanthraniloyl) nucleotides bound to the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             METHYLATION (DI-) (POTENTIAL).
ALKYLATION (SH-1).
PHOSPHORYLATION (BY MHCK).
PHOSPHORYLATION (BY MHCK).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COILED COIL (POTENTIAL). ATP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACTIN-BINDING. ACTIN-BINDING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR000048; -.
InterPro; IPR001609; -.
Pfam; PF00612; IO; 1.
Pfam; PF00063; myosin_head; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; M14628; AAA33227.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PDB; 1LVK; 28-JAN-98.
DictyDb; DD01008; mhcA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          752
130
678
1823
1833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1MMA; 03-DEC-97
1MMD; 17-AUG-96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           17-AUG-96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  03-DEC-97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   03-DEC-97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         S00250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A26655; A26655
S00250; S00250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LMND;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1VOM;
1LVK;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MWN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MNE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MMG;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
DOMAIN
DOMAIN
NP_BIND
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOD_RES
MOD_RES
MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR;
PIR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PDB;
PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PDB;
```

```
S48385; S48385.
S0001411; MLP2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              YIO9_YEAST
P40457;
   SITE
SITE
SEQUENCE
                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 13
YIO9_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIR;
SGD;
      FT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DDT THE BANK THE BANK
                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                           12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JQ
                                                                                                                                                                                                                                                                                                 1673 IKSLVAEVDEVKEQLEDEI------LAKDKLVKAKRALEVELEEVRDQLEEEEDSRS 1723
                                                                                                                                                                                                                                                                                                                                                                                                                    1772 EDEKKKLNES-----ERAKKRLESENEDFLAKLDAEVKNRSRAEKDRKKYEKDLKD 1822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1877 DNLRAQIEDEGKIKMRLEKEKRALEGELEEL----RETVEEAEDSKS-----EAEQS 1924
                                                                                                                                                                                                                                                                                                                                                                                   81 IIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGISYYEKVLAKYKDDLESIKKVI 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          141 KEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV-----NKIDDYLIN 194
                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                   22 LKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKYFLDV-LESDLMQFKHISSNEY 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cell 69:1213-1226(1992).
-!-FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS 'HUMAN RETICULOCYTE CELLS.
-!-SUBUNIT: HOMODIMER (POPENTIAL).
-!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKKLINDDTKKDMLGKL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    255 LSTGLVQNFPNTIISKL-----IEGKFQDMLNISQHQCVKKQCPENSGCFRHLDEREEC
                                                                                                                                                                                           72;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RETICULOCYTE BINDING PROTEIN 1. EXTRACELLULAR.
                                                                                                                            Length 2116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBL_TaxID=31273;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-92315338; PubMed-1617731;
Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
"A reticulocyte-binding protein complex of Plasmodium vivax
                                                                                                                                                                                        63; Mismatches 132; Indels
29 PHOSPHORYLATION (BY MHCK).
243871 MW; 2FC3770BB1EE56A1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     309 KCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSN 350
                                                                                                                            DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
RETICULOCYTE BINDING PROTEIN 1 PRECURSOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2869 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Receptor; Signal; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
CYTOPLASMIC.
                                                                                                                         6.6%; Score 137;
21.9%; Pred. No. 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Plasmodium vivax (strain Belem)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M88097; AAA29743.1; -. HSSP; P36956; 1AM9.
                                                                                                                                             Best Local Similaring
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
2029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2869
2807
2826
2869
2029 202
2116 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2808
2827
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                merozoites."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RBP1_PLAVB
Q00798;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
DOMAIN
                             SEQUENCE
                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Malaria;
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPI_PLAVE

ID RBPI_PLAVE

ID RBPI_PLAVE

DT 01-APR:

DT 01-APR:

DE RETICUL

GN RBPI.

OX RCBI_TI

RP SEQUEN

RX MEDLIN

RX MEDLIN

RX MEDLIN

RY A ret.

RI Mercoco

CC -!- FU

CC -!- SU

CC -!- SU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          195
E S
                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                      17;
                                                                                                                                                                                                                                                                                                                                                                                           : |:| :: |: | : | : | : | : | 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       122 YEKVLAKYKDDLESIKK ------VIKEEKERFPSSPPTTPPSPAKTDEQKKESKFLPFL 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 77 SNEYIIEDSFKLLNSEQKNTLLKSYKYIK------ESVENDIKFAQEGI----SY 121
                                                                                                                                                                                                                                                                  Indels 81; Gaps
                                                                                                                                                                                                                                                                                                                                                34 KQIEKNIFTFNLNLNDI------LNSRLKKRKYFLDVL----ESDLMQFKHIS 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Parall B.G. Badcock K., Bankier A.T., Bowman S., Brown D., Barrell B.G., Badcock K., Bankier A.T., Dear S., Devlin K., Fraser A., Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Jones M., Louis E., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M., Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D., Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V., Walsh S.V., Whitehead S., Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2384 YSVIEAEVK-TLEEIDRDYGDNYQIVEEHKKQFSILIDRTNALMDDIEIFKKENNYNLME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       175 TNIETLYNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDD-----KI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            227 DLFKNHNDFEAIKK----LINDDTKKDMLGKLLSTG-----LVQNFPNTIISKLIEGKFQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         278 DMLNISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQ--EGDKCVENPNPTCNENNGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
01-FEB-1995 (Rel. 31, Last annotation update)
HYPOTHETICAL 195.1 KDA PROTEIN IN DNA43-UBI1 INTERGENIC REGION.
                                                                                                                                                                                            Length 2869;
1030 1032 CELL ATTACHMENT SITE (POTENTIAL) 2599 2601 CELL ATTACHMENT SITE (POTENTIAL) 2869 AA; 330213 MW; B9DBE442205EBCFF CRC64;
                                                                                                                                                                                        6.6%; Score 136.5; DB 1;
21.7%; Pred. No. 8.8;
tive 70; Mismatches 145;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 1679 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2602 MDAIFSQMSADRNPNEYK 2619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; 238059; CAA86129.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          336 CDADAKCTEEDSGSNGKK 353
                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                               Best_Local Similarity
Matches 82; Conserv
```

```
Best Local Similarity
Matches 75; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
     17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUBDINIT: MULTIMERIC. COMPOSED OF VARYING SIZED, DISULFIDE-LINKED MULTIMERS, THE SMALLEST OF WHICH IS A HOMOTRIMER. PROTECLYSIS OF THE PROMULTIMERIN IN THE N-TERMINAL REGION, LEADS TO THE MATURE PISS FORM THAT IS STORED IN PLATELETS.

TISSUE SPECIFICITY: SYNTHESIZED BY ENDOTHELIAL CELLS AND MEGARARYOCYTES. STORED IN PLATELET ALPHA GRANULES AND ENDOTHELIAL CELL WEIBBL-PALADE BODIES, FOLLOWING ACTIVATION OF THESE CELLS, IT IS RELEASED AND ATTACHED TO MEGARARYOCYTES, PLATELETS, ENDOTHELIUM AND SUBENDOTHELLUM OF BLOOD VESSELS. NOT FOUND IN PLASMA FORIND IN
                                                                                                                                                                                                                               1079 SVLIEKVDDTAANNGDKDHLK-LVSLFSNLRHERNSLETKLTTCKREL----AFVKQKN 1132
                                                                                                                                                                                                                                                                                                                          1193 KNREIYKQLNDRQEEISRLQRDLIQTKEQVSINSNKILVYESEMEQCKQRYQDLSQQQKD 1252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1253 AOKKD-----IEKLTNEISD----LKGKLSSAENANADLENKFNRLKKQAHE 1295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1296 KLDASKKQQAALTNELNELKAIKDKLEQDLHFEN-----AKVIDLDTKL-KAHELQSE 1347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1348 DVSRDHEKDTYRTLME----EIESLKKELQIFKTANSSSDAFE-----KLKVNMEKE 1395
                                                                                                                                                                                                                                                                                              60 YFLDVLESDLMQFKHISSNEY----IIEDSFKLLNSE--QKNTLLKSYKYIKESVEN--- 110
                                                                                                                                                                                                                                                                                                                                                                                               -----DIKFAQEGISYYEKVLAKYKDDLE-SIKKVIKEEKEKFPSSPPTTPPSPAKTD 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Brown C., Kelton J.G.;
"The CDNA sequence of human endothelial cell multimerin. A unique protein with RGDS, coiled-coil, and epidermal growth factor-like domains and a carboxyl terminus similar to the globular domain of complement C1q and collagens type VII and X.";
J. Blol. Chem. 270:18246-18251(1995).
-!- FUNCTION: CARRIER PROTEIN FOR PLATELET (BUT NOT PLASMA) FACTOR V/VA. MAY PLAY A ROLE IN THE STORAGE AND STABILIZATION OF FACTOR IN PLATELETS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          163 EQKKESKFLPFLTNIETLYNNLVNKIDDYLINLKAKI------NDCNVEKDEAHV 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          212 KI------TKLSDLKAIDDKI--DL-FKNHNDFEAIKKLINDDTKKDMLGKLLST 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        258 GLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDEREECKCLLNYKQE 317
                                                                                                                                                   Gaps
                                                                                                                                                                                                3 SVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQ---IEKNIFTFNLNLNDILNSRLKKRK 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                   88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A., AND SEQUENCE OF 368-376.
TISSUE-Endothelial cells;
MEDLINE-65355440; PubMed-7629143;
Hayward C.P.M., Hassell J.A., Denomme G.A., Rachubinski R.A.,
                                                                                                  DB 1; Length 1679;
                                                                                                6.6%; Score 136; DB 1; Length 16
22.4%; Pred. No. 5.2;
tive 69; Mismatches 127; Indels
                         1679 AA; 195141 MW; 298950CC52202D8F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (Rel. 36, Created)
(Rel. 36, Last sequence update)
(Rel. 39, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 1228 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ENDOTHELIAL CELL MULTIMERIN PRECURSOR.
                                                                                                                                                   82; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens (Human).
Hypothetical protein.
                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1396 KDRIID 1401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         318 GDKCVE 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15-JUL-1998 (
15-JUL-1998 (
30-MAY-2000 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ECM_HUMAN Q13201;
                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 14
ECM_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                               111
                                                                                                                                                   Matches
                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
  SS
                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DDATE DDATE
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
              -:- PTM: THE N-TERMINUS IS BLOCKED.
-:- PTM: XTENSTYELY NGLYCOSTLATED.
-:- DISBASE: DEFICIENCY IN MULTIMERIN DUE TO PROTEOLYTIC DEGRADATION WITHIN THE PLATELET ALPHA GRANULES IS ASSOCIATED WITH AN AUTOSOMAL DOMINANT BLEEDING DISORDER (FACTOR V QUEBEC).
-:- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLY-GLN. CELL ATTACHMENT SITE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ENDOTHELIAL CELL MULTIMERIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EAA98B0A17E2C4CD CRC64;
VASCULAR TISSUES SUCH AS PLACENTA, LUNG, AND LIVER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COILED COIL (POTENTIAL)
COILED COIL (POTENTIAL)
COILED COIL (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COILED COIL (POTENTIAL)
COILED COIL (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COIL (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      domain; Coiled coil.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EGF-LIKE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COILED
                                                                                                                                                                                                                                                                                                                                                                                                                        PROSITE; PS01113; COLO; IL PROSITE; PS01113; CIO; IL PROSITE; PS01186; EGF_1; I. Signal; Glycoprotein; EGF-11ke sIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                               COMPLEMNTC1Q.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AA; 138071
                                                                                                                                                                                                                                                                                                        EMBL; U27109; AAC52065.1;
                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR000561; -.
InterPro; IPR001073; -.
Pfam; PF00008; EGF; 1.
PRINTS; PR00007; COMPLEM
                                                                                                                                                                                                                                                                                                                                         InterPro; IPR000152; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           056
065
076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            981
1020
1075
1228
                                                                                                                                                                                                                                                                                                                          MIM; 601456;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
```

15;

Gaps

:99

Indels

6.5%; Score 135; DB 1; Length 1228; 22.3%; Pred. No. 4.2;

63; Mismatches 132;

Conservative

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation between the Dinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL. SIMILARITY: BELONGS TO THE P115(TAP)/USO1/YBL047C FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bai Y., Symington L.S.; Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases. 
-!- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
                                                   GLKSKSINVLIRD--IVREQFKIFQNDM---QETVAQLFKTVSSLSEDLESTRQIIQKVN 431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EKPPSSPPTTPPSPAKTDEQKKESKFL-----PFLTNIETLYNNLVNKIDDYLINLKAK 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                   199 INDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHND----FEAIKKLINDDTKKDMLGKL 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDEREECKCLLNY 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUBCELLUIAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE ER AND THE GOLGI COMPLEX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---CEKPIKELEVKQTHLEGA------LEQEHSRSILYYESLNKTLS--KLKEVHEQL
29 YRSLK-KQIEKNIFTFNLNLNDILNSRLKKRKYFLDVLESDLMQFKHISSNEYIIEDSFK
                                                                                                                                                     LLNSEQKNTLLKSYKYIKES---VENDIKFAQEGISYYEKVLAKYKDDLESIKKVIKEEK
                                                                                                                                                                                                                                                                                                                                                                        ESVVSI-----AAQQKFVLVQENRPTLTDIVELRNHIVNVRQEMTLT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "A cytoskeleton-related gene, usol, is required for intracellular protein transport in Saccharomyces cerevisiae."; J. Cell Biol. 113:245-260(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                USOI OR INTI OR YDL058W.
Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetales; Saccharomycetales
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 782-1790 FROM N.A.
Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Submitted (FEB-1993) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSN 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            580 SLEMEK -- ESLRGECEDMLSKCRNDFKFOLKDTEEN 613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAY-1992 (Rel. 22, Created)
01-MAY-1992 (Rel. 22, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
INTRACELLULAR PROTEIN TRANSPORT PROTEIN USOL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 1790 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=91185402; PubMed=2010462;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 1-8 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN=X2180-1A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Kendrick K.E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    USO1_YEAST
P25386:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPLEX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /amasaki M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 15
USO1_YEAST
                                                                                                                                                                                                                                                                                                  145
                                                                                                                                                 88
                                                                                                                                                                                                                                                                                                                                                                        432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     315
                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A PART OF THE PROPERTY OF THE
```

```
18;
                                                                                                                                                                                                                                                                                                                                                                                                                                             1235 LKKKNETNEASLLESIKSVESETVKIKELQDECNFKEKEVSE--LED--KLKASEDKNSK 1290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1179 ISQLNDEITS-TQQENESIKKKNDELEGEVKAMKSTSEEQ---SNLKKSEIDALNLQIKE 1234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -EKFPSSPPTTPPSPAKTDEQKKESKFL-PFLTNIETLYNNLVNKIDDYLL-----NLK 196
                                                                                           fransport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil;
                                                                                                                                                                                                                                                                                                                                                    Indels 109; Gaps
                                                                                                                           COILED COIL (POTENTIAL).
CHARGED (HYPER-HYDROPHILIC).
DISPENSABLE FOR THE PROTEIN FUNCTION.
                                                                                                                                                                                                                                                                                                                                                                          ISVTMDNILSGFENEYDVIYLK - - PLAGVYRSLKKQIEKNIFTFNLNLNDI - - LNSRLKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AKINDCNVEKDEAHVKIT-----KLSDLKAIDDKIDLFK-------NHN
                                                                                                                                                                                                                                                                                                                                                                                                                       58 RKYFLDVLESDLMQ------FRHISSNEYIIEDSFKLLNSEQKNT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---LLKSYKYIKESVE---NDIKFAQEGISYYEKVLAKYKDDLESIKKVIKEEK----
                                                                                                                                                                                                                                                                                                                              Length 1790;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6CE2B216E9FD4818 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ---IKKLINDDTK-----KDMLGKLLST
                                                                                                                                                                                                                                                                                  REF.
                                                                                                                                                               ASP/GLU-RICH (ACIDIC) POTENTIAL.
                                                                                                                                                                                                                                                                                                                           Query Match
6.3%; Score 130.5; DB 1;
Best Local Similarity 24.3%; Pred. No. 11;
Matches 85; Conservative 50; Mismatches 106;
                                                                                                                                                                                                                                                                                  NI)
                                                                                                                                                                                  G -> E (IN REF.

E -> K (IN REF.

I -> V (IN REF.

N -> S (IN REF.

G -> S (IN REF.

G -> S (IN REF.

I -> V (IN REF.

I -> V (IN REF.

B -> S (IN REF.

D -> DEEDDEE (IN
                                                                                                                  GLOBULAR HEAD.
send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                            ₩;
       EMBL; X54378; CAA38253.1; -...
EMBL; L03188; AAB00143.1; -...
                                                                                                                                                                                                                                                            ,60.
1661
1772
206424 P
                                  L03188; AAB00143.1;
U53668; AAB66659.1;
                                                                                                                                                                                                                                                                               1772 177
1790 AA;
                                                         PIR; A38455; A38455.
                                                                               SGD; S0002216; USO1
                                                                     P80220; 1DIP
                                                                                                   Calcium-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DFEA--
                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                        CA_BIND
CONFLICT
CONFLICT
                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                        DOMAIN
                                                                                                                                                   DOMAIN
                                                                                                                                                                DOMAIN
                                                                                                                 DOMAIN
                                                                                                                             DOMAIN
                                                                    HSSP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  145
                                                                                                                                                                                                                                                                                                                                                                           ~
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              234
g
                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

Search completed: August 8, 2001, 12:32:29 Job time: 197 sec

THIS PAGE BLANK (USPTO)

Ŧ

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein – protein search, using sw model

Run on:

August 8, 2001, 12:30:49; Search time 43.63 Seconds (without alignments) 687.893 Million cell updates/sec

US-09-500-376-3 2052 1 AISVTMDNILSGFENEYDVI......SNFLGISFLLILMLILYSFI 394 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

219241 seqs, 76174552 residues Searched:

Total number of hits satisfying chosen parameters:

219241

Minimum DB seq length: 0 Maximum DB seq length: 200000000

o

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR\_68:\* Database :

pirl:\* pir2:\* pir3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Description		major merozoite su	major merozoite su	major merozoite su	merozoite surface		merozoite		major merozoite su	probable major sur	ro	major blood-stage				rhoptry protein -	hypothetical prote	യ	hypothetical prote		pr	Ψ	hypothetical prote	>	hypothetical prote	Н	¥	hypothetical prote	
A		A26868	A54498	SAZOGM	S47282	A45948	A45545	S05603	SAZQK1	A24594	A39401	A45604	A28121	A45532	A45546	T28676	E71606	T18372	G70163	A64505	T18499	T28677	T18355	T28317	T34081	S48385	E81339	C71618	10000
DB	:	7	7	Н	7	~	~	7	۲	~	7	7	7	7	7	7	~	~	7	~	~	7	7	7	7	7	7	7	c
Length		1701	1701	1726	651	1726	400	1639	1631	1640	1726	1751	680	1772	1785	2401	980	1939	2166	1169	1619	2269	1187	1127	1130	1679	1250	1712	,,,,
% Query Match		100.0	99.1	97.8	6.96	8.96	61.1	61.1	S	53.2	40.3	•	28.9	ω.	27.7	8.2	8.1		7.6	7.5	7.5		7.3		7.2		7.1	7.1	,
Score	1 0	2022	2033	2006	1987.5	1987	1254.5	1254.5	1133	1091	826	815	593.5	593.5	269	168.5	166	156.5	155.5	153.5	153.5	152	150	149.5	147	147	146.5	145.5	
Result No.		-	۲۹.	m	4	S	9	7	æ	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	ć

qq δ

conserved hypothet	hypothetical prote	surface membrane p	hypothetical prote	Plls protein homol	conserved hypothet	protein g377 - mal	hypothetical prote	hypothetical prote	membrane protein p	hypothetical prote	cag island protein				
						. :									
F75103	465	522	501	502	545	003	087	105	387	114	203	102	356	573	328
F7	A64	C71	T18	H64	E64	PC6	S67	E70	A70	T184	T185	B593	T183	F966	A71
2 F7	2 A64	2 C71	2 T18	2 H64	2 E64	2 PC6	2 S67	2 E70	2 A70	2 T184	2 T185	2 B593	2 T183	2 F966	2 A71
880 2 F7	7	ď	~	7	7	7	7	7	7	~	7	7	~	7	7
880 2	1005 2	1979 2	3394 2	671 2	793 2	624 2	821 2	819 2	978 2	~	442 2	652 2	1079 2	1313 2	7
880 2	6.8 1005 2	6.8 1979 2	6.8 3394 2	6.7 671 2	6.7 793 2	6.7 624 2	6.7 821 2	6.7 819 2	6.6 978 2	6.6 3119 2	6.6 442 2	6.6 652 2	6.6 1079 2	6.6 1313 2	6.6 1819 2

## ALIGNMENTS

RESULT 1 A26868 merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) C;Species: Plasmodium falciparum C;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 09-Jun-2000 C;Accession: A26868 C;Accession: A26868 Mill: 195, 273-287, 1987 A;Title: Allelic dimorphism in a surface antigen gene of the malaria parasite Plasmod A;Reference number: A26868; MulD:88011243 A;Reference number: A26868 A;Molecule type: DNA A;Residues: 1-1701 - Transport merozoite surface antigen C;Superfamily: major merozoite surface antigen C;Superfamily: major merozoite surface antigen E;1-19/Domain: signal sequence #status predicted <nat> F;20-1701/Product: major merozoite surface antigen #status predicted <nat></nat></nat>	Query Match Best Local Similarity 100.0%; Pred. No. 9.1e-101; Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNITFFNLALNDILNSRLKKRKY 60	61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120 	1 YYEKVLAKYKDDLESIKKVIKBEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180 	1 INNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK 240 	1 LINDDFKKDMLGKLLSFGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300	1 HLDEREECKCLLNYKOEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 360 
RESULT A26868 major me C. Specie C. Date: C. Access R. Tanabe A. Refere A. Refere A. Residu C. Superf. C. Superf. C. Superf. C. Superf. C. Superf.	Query Match Best Local Matches 39	130	13	121	181	241	301
PANTA	O H Z	Qy	Qy	Oy Dp	Qy Db	QY Db	Qy Db
				- · · · · - · - · · · · · · · · · · · ·			

a

ant

carb

ö

Query Match 99.1%; Best Local Similarity 99.2%; Matches 391; Conservative

61

ò

qq

g ò

A; Molecule type: DNA

RESULT

```
A.Cross-references: EMBL.X03831
C.Comment: The merozoite stages of different strains have strain-specific surface a.C.Comment: The merozoite stages: sporozoite, merozoite, and gametocyte. (C.Comment: P. falciparum has three stages: sporozoite, merozoite, and gametocyte. (S.C.Superfamily: major merozoite surface antigen; tandem repeat
F.1-19/Domain: signal sequence #status predicted <SIG>F.20-1726/Product: major merozoite, surface antigen #status predicted <MAT>F.70-87,91-96,100-105,109-120/Region: 3-residue repeats (S-G-T)
F.77-765/Region: 3-residue repeats (T-E-E)
F.133,272,501,567,638,827,839,924,944,990,1016,1114,1221,1613,1658/Binding site: c.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C. Species: Plasmodium falliparum
A. Variety: strain R0-71
C. Species: Plasmodium falliparum
A. Variety: strain R0-71
C. Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 09-Jun-2000
C. Accession: S47282
R. Tolle, R.; Bujard, H.; Cooper, J.A.
submitted to the EMMB. Data Library, July 1994
A. Description: Plasmodium falciparum: recombination within the C-terminal region
A. Reference number: S47282
A. Molecule type: DNA
A. Residues: 1-651 < TOL>
A. Reperimental source: strain R0-71
C. Superfamily: major merozoite surface antigen
C. Keywords: glycoprotein; merozoite; surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   merozoite surface antigen 1 - malaria parasite (Plasmodium falciparum) (strain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1393 FLDVLESDLMOFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 1452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1453 YYEKVLAKYNDLESIKKVIKEEKEKPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 1512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 AISVIMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEOKKESKFLPFLTNIETL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1;
                                                                                                                                                                                                                                                                                                                                                                                                 ő
                                                                                                                                                                                                                                                                                                                                             Length 1726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length
                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 1987.5; DB 2;
Pred. No. 7.8e-98;
2: Mismatches 7;
                                                                                                                                                                                                                                                                                                                                                                                                 7;
                                                                                                                                                                                                                                                                                                                                          Ouery Match 97.8%; Score 2006; DB 1; Best Local Similarity 98.0%; Pred. No. 2.5e-98; Matches 386; Conservative 1; Mismatches 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               tch 96.9%; al Similarity 97.5%; 384; Conservative
Residues: 1104-1726 <WEB2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 384; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ч
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
S47282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         . 업
     Add 498
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (is C; Species: Plasmodium falciparum
C; Species: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C; Accession. A54498
R; Peterson, M.G.; Coppel, R.L.; McIntyre, P.; Langford, C.J.; Woodrow, G.; Brown, G.V.; Mol. Blochem. Parasitol. 27, 291-302, 1988
A; Title: Variation in the precursor to the major merozoite surface antigens of Plasmodiu A; Reference number: A54498
A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ğ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          malaria parasite (Plasmodium falciparum) (st
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (s N;Alternate names: 195K glycoprotein C;Species: Plasmodium falciparum c;Species: Plasmodium falciparum (c;Species: Plasmodium falciparum (c;Date: 30-Sep-1987 #sequence_revision 31-Mar-1991 #text_change 09-Jun-2000 (c;Accession: A23386; S06361 (S;Weber, J.L.; Leininger, Ww., Lyon, J.A. Nucleic Acids Res. 14, 3311-332, 1986 (A;Title: Variation in the gene encoding a major merozoite surface antigen of the human A;Reference number: A23386; MUID:86205236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Title: Merozoite surface protein sequence from the Camp strain of the human malaria A;Reference number: S06361; MUID:88143999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1428 YYEKVLAKYKODLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 1487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 YYEKVLAKYKDDLESIKKVIKEEKERFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                        A; Residues: 1-1701 <PET>
A; Residues: 1-1701 <PET>
A; Cross-references: GB:M19143; NID:g160412; PIDN:AAA29653.1; PID:g160413
C; Superfamily: major merozoite surface antigen
C; Keywords: Surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 301 HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 1701;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 2033; DB 2;
Pred. No. 9.1e-100;
0; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1668 PDSYPLEDGIFCSSSNFLGISFLILMLILYSFI 1701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
```

181

ð

QQ

Q

ò

241

ò

QQ

qq

ò

a

δ

E

of

÷

A; Accession: A23386
A; Molecule type: DNA
A; Residues: 1-1104 <WEB1>
A; Cross-references: EMBL: X03831
A; Weber, J.L.; Sim, B.K.L.; Lyon, J.A.; Wolff, Nucleic Acids Res. 16, 1206, 1988

type: DNA

Accession: S06361 Molecule type: DN

RO-71)

3

"

음

g

ò

g

ŏ

පු

ò

g

ò

셤

ò

```
- malaria parasite (Plasmodium falciparum) (fragments
                                                                                     major merozoite surface antigen - malaria parasite (Plasmodium falciparum) (fragmen C; Species: Plasmodium falciparum
C; Species: Plasmodium falciparum
C; Date: 17-Feb-1994 **sequence_revision 17-Feb-1994 **text_change 09-Jun-2000
C; Accession: A45545
R; Blackman, M.J; Ling, I.T.; Nicholls, S.C.; Holder, A.A.
Nol: Biochem. Parasitol. 49, 29-33, 1991
A;Title: Proteolytic processing of the Plasmodium falciparum merozoite surface prote A; Reference number: A45545; MUID:92131048
A; Reference number: A45545
A; Molecule type: DNA
A; Molecule type: DNA
A; Residues: 1-400 < BLA>
A; Nolecule type: DNA
C; Superfamily: major merozoite surface antigen
C; Superfamily: major merozoite surface antigen
C; Keywords: qlvconrotoir.
                                                                                                                                                                                                                                                 protei
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             parasite (Plasmodium falciparum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ä,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 major merozoite surface antigen precursor - malaria parasite (Plasmodium fall N.Alternate names: gp195 surface antigen C.Species: Plasmodium falciparum C.Species: Plasmodium falciparum C.Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jun-2000 C.Accession: S05603; S04850  
R.Wyler, P.J. submitted to the EMBL Data Library, April 1989  
A.Reference number: S05603  
A.Accession: BNBL:XIS063; NID:99896; PIDN:CAA33163.1; PID:99897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       151 SEKYKSDLDSIKKYI--------NDKQGENEKYLPFLNNIETLYKTVN 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TKKDMLGKLLSTGLVQIFPN---TIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHL 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31 IDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRENFKNVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DERECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 400;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 61.1%; Score 1254.5; DB 2; Best Local Similarity 61.2%; Pred. No. 2.1e-59; Matches 240; Conservative 54; Mismatches 73;
SYPLEDGIFCSSSNFEGISFLILMLILYSFI 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SYPLFDGIFCSSSNFLGISFLLILMLILYSFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         363
                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1333 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRY 1392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1633 HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 1692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 360
                                     FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                   FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                           YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                 437
                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                               HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                               LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 1987; DB 2;
Pred. No. 2.5e-97;
1; Mismatches 10;
                                                                                                                                                                                                                                                                                                                                                                                                                       361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 96.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 383; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                     618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1573
                 61
                                                                                         121
                                                                                                                                                                                                                                                                                   498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         241
                                                                                                                                                                                                                                             241
                                                                                                                                                                                                                                                                                                                       301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       301
```

ò g g

ò

g

ò

g ŏ q

ò

Nucleic Acids Res. 17, 5401, 1989

```
Cincession: A24594

Sith of the control of the precursor to the three major surface antigens of Pl
Airtle: Primary structure of the precursor to the three major surface antigens of Pl
Airtle: Primary structure of the precursor to the three major surface antigens of Pl
Airtle: Primary structure of the precursor to the three major surface antigens of Pl
Airtle: Primary structure of the precursor to the three major surface antigens of Pl
Airtle: Primary structure of the precursor to the three major surface antigen
Cisteywords: Surface antigen
                                                  probable major surface antigen (83K, 19K, 42K) precursor - malaria parasite
C;Species: Plasmodium falciparum
C;Date: 29-Aug-1987 #sequence_revision 29-Aug-1987 #text_change 09-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1330 ESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVLGYYKIL 1389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           246 TKKDMLGKLLSTGLVQIFPN---TIISKLIEGKF-QDMLNISQHQCVKKQCPENSGCFRH 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ESDLMQFKHISSNEY11EDSFKLLNSEQKNILLKSYKY1KESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LDEREECKCLLNYKOEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSKKKITCECTKP 361
                                                                                                                                                                                                                                                                                                                                                                -----NDKQGENEKYLPFLNNIETLYKTVN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNLNDILNSRLKKRKYFLLDVL
              MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVL
                                                                                                                                                                                                                                                                                              LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKP
                                                                                                                                                                                                                                                                                                                                                                                                                                     NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEOKKESKFLPFLTNIETLYNNLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TKKDMLGKLLSTGLVQIFPN---TIISKLIEGKF-QDMLNISQHQCVKKQCPENSGCFRH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Pred. No. 4.2e-50; 60; Mismatches 90;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            53.2%; Score 1091; DB 2; 55.2%; Pred. No. 4.2e-50;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity 55.2%
es 217; Conservative
                                                                                                                                                                                                                                                                                                                                                                1381 SEKYKSDLDSIKKYI----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  σ
                                                                                                                                                                                                                                                                                              126
                                                                                                                                                    99
                                                                                                                                                                                                                                                                                                                                                                                                                                     186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                            g
                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                           δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Residues: 1-1631 <AMC>
C; Comment: The merozoite stages of different strains have strain-specific surface antige C; Comment: The merozoite stages: sporozoite, merozoite, and gametocyte. The me C; Superfamily: major merozoite surface antigen C; Superfamily: major merozoite; surface antigen c; Superfamily: major merozoite; surface antigen; trandem repeat; transmembrane prote F; 1-19/Domain: signal sequence #status predicted <SIG>
F; 7-19/Domain: signal sequence #status predicted <SIG>
F; 67-84/Region: 3-residue repeats (S-G-T/P)
F; 1614-1631/Domain: membrane anchor #status predicted <WAT>
F; 1614-1631/Domain: membrane anchor #status predicted <WAT>
          sequence of the gp195 (MSA-1) gene from Plasm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        fald
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (st
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SAZOK1
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (C;Species: Plasmodium falciparum c;Species: Plasmodium falciparum C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jun-2000 C;Accession: A25120 C;Accession: A25120 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Title: Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) ge A; Reference number: S04850; MUID:89345116
A; Accession: S04850
A; Accession: S04850
A; Residues: 1504-1639 < MYL2>
A; Residues: 1504-1639 < MYL2>
A; Cross-references: EMBL:X15063
C; Superfamily: major merozoite; Surface antigen
C; Superfamily: major merozoite; surface antigen
F; 1-19/Poomain: signal sequence #status predicted <SIG>F; 20-1639/Product: major merozoite surface antigen #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              186 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPD 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    246 TKKDMLGKLLSTGLVQIFPN -- - TIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25;
                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 1631;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  55.2%; Score 1133; DB 1; 57.0%; Pred. No. 2.5e-52;
                                                                                                                                                                                                                                                                                                                                                                                                                                 61.1%; Score 1254.5; Di
61.2%; Pred. No. 1e-58;
tive 54; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 61.28
Matches 240; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
```

303

셤

ò

99

ð g ò g ò g ò Q ò

a

ò

4

Length 1640; Indels LDEREECKCLLNYKOEGSKCVENSNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKP

1548

Q

4;

Gaps

26;

Indels

86;

57; Mismatches

224; Conservative

Matches

S

9

Gaps

```
R;Burns Jr., J.M.; Daly, T.M.; Vaidya, A.B.; Long, C.A.
Proc., Natl. Acad. Sci. U.S.A. 85, 602-606, 1988
A;Title: The 3' portion of the gene for a Plasmodium yoelii merozoite surface antigen
A;Reference number: A28121; MUID:88124889
                                                                                                                                                                                                                                                                                                                                              ----KK--NEVKSSGLLEKLMKSKLIKENESKEILSQLLNVQTQLLTMSSEHTCIDTNV 1652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C.Species: Plasmodium yoelii
C.Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 26-Aug-1999
                                                                                                                                                                                                                                             1491 GVKTEIKKVEDDI------KKQDEELKKLGNVNSQDSKKNEFIAKKAELEKYLPF 1539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            174 LTNIETLYNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTN 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       293 PENSGCERHLDEREECKCLLNYKOEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRK 352
                                                                                                                                                                                                                                                                                                                   75 ISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKVLAKYKDDLE 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          369 ELDLFKDLSTNKYVIRNPYQLLDNDKKDKQIŲNLKYATKGINEDIETTTDGIKFFNKMVE 428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKVLA 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          128 KYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKE--SKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15 NEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVLESDLMQFKH 74
A;Note: sequence extracted from NCBI backbone (NCBIN:83591, NCBIP:83592)
C;Superfamily: major merozoite surface antigen
C;Reywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           234 DFEAIKKLINDDTKKDMLGKLLSTGLV-QIFPNTIISKLIEGKFQDMLNISQHQCVKKQC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVLES
                                                                                                                                                                    44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Leu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 680;
                                                                                                                       Length 1751;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         major merozoite surface antigen - Plasmodium yoelii (fragment)
                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A)Cross-references: GB:J03612; NID:g160678; PID:g160679
A)Experimental source: strain 17XL
A)Note: the authors translated the codon GTA for residue 429
C; Superfamily: major merozoite surface antigen
C; Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28.9%; Score 593.5; DB 2; 32.5%; Pred. No. 2.7e-24; ive 96; Mismatches 137;
                                                                                                                     Ouery Match 39.7%; Score 815; DB 2; L Best Local Similarity 41.0%; Pred. No. 1.6e-35; Matches 163; Conservative 82; Mismatches 109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             353 KITCECTKPDSYPLFDGIFCSSSNFLGISFLLILMLIL 390
                                                                                                                                                                                                                                                                                                                                                                                                               135 SIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Accession: A28121
A; Molecule type: DNA
A; Residues: 1-680 <BUR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity
Matches 128; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; Accession: A28121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Φ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           89
                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         윤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                              merozoite surface antigen 1 precursor - Plasmodium vivax
C;Species: Plasmodium vivax
C;Accession: A3401
R;del Portillo, H.A.; Longacre, S.; Khouri, E.; David, P.H.
R;del Portillo, H.A.; Longacre, S.; Khouri, E.; David, P.H.
R;del Portillo, H.A.; Longacre, S.; Khouri, E.; David, P.H.
A;Title: Primary structure of the merozoite surface antigen 1 of Plasmodium vivax reveal
A;Reference number: A3401
A;Reference number: A3401
A;Residues: DIA
A;Residues: 1-1726 cDEL>
A;Residues: 1-1726 cDEL>
A;Residues: 1-1726 cDEL>
A;Cross-references: GB:M60807
C;Superfamily: major merozoite surface antigen
C;Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AA5604

AA561

AA5604

C; Species: Plasmodium vivax
C; Species: Plasmodium vivax
C; Species: Plasmodium vivax
C; Species: Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
C; Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
C; Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
R; Cibson, H.L.; Tucker, J.E.; Kaslow, D.C.; Krettli, A.U.; Collins, W.E.; Kiefer, M.C.; Mol. Blochem. Parasitol. 50, 325-333, 1992
A; Title: Structure and expression of the gene for Pv200, a major blood-stage surface ant A; Reference number: A45604
A; Reference number: A45604
A; Accession: A45604
A; Abolecule type: DNA
A; Residues: 1-1751 <GIB>A; Residues: 1-1751 <GIB>A; Cross-references: GB:M75674; NID:g160608; PID:g457336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1579 VKSSGLLEKLMKSKLIKENESKEILSQLLNVQTQLLTMSSEHTCIDTNVPDNAACYRYLD 1638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKVLAKYKDDL- 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---ESIKKV---IKEEKEKFPSSPPTTPPSPAKTDEQKKES---KFLPFLTNIETLYNNL 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DTKKDMLGKLLSTGLV-QIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLD 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVLESDLMQFKH 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           304 EREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; Score 826; DB 2; Length 17; Pred. No. 4.2e-36; 79; Mismatches 119; Indels
                 |||:||:||||||:|| ||:::| |
EPLFEGVFCSSSFLSLSFLLLMLLFL 1721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          364 YPLFDGIFCSSSNFLGISFLLILMLIL 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 40.3%;
Best Local Similarity 43.2%;
Matches 167; Conservative 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            245
                      ò
                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
```

œ ;

Gaps 67 9

```
mijor merozoite surface antigen precursor - Plasmodium chabaudi chabaudi (Species: Plasmodium chabaudi di chabaudi chabaudi di chacesion: A45546 R;Deleersnijder, W.; Hendrix, D.; Bendahman, N.; Hanegreefs, J.; Brijs, L.; Hamers-Ca A, Rochesnijder, W.; Hendrix, D.; Bendahman, N.; Hanegreefs, J.; Brijs, L.; Hamers-Ca A, Reference number: A45546; MUID:91218805 A, Reference number: A45546; MUID:91218805 A, Rochesion: A45546; MUID:91218805 A, Residues: 1-1785 CDEL>
A, Residues: 1-1785 CDEL>
A, Residues: 1-1785 CDEL>
A, Residues: 1-1785 CDEL>
A, Cossereferences: GB:M34947; NID:g160597; PID:g160598 C, Superfamily: major merozoite surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A,Accession: T28676
A,Status: preliminary: translated from GB/EMBL/DDBJ
A,Molecule type: DNA
A,Residues: 1-2401 <SIN>
A,Residues: 1-2401 <SIN>
A,Cross-references: EMBL:U36927; NID:91041784; PID:91041785; PIDN:AAB41263.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             C;Species: Plasmodium yoelii
C;Date: 15-0ct-1999 #sequence_revision 15-0ct-1999 #text_change 01-Dec-2000
C;Accession: T28676, A45521.
R;Sinha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.
Mol. Biochem, Parasitol. 76, 329-332, 1996
A;Fitle: Comparison of two members of a multigene family coding for high-mol
A;Reference number: 220507; MUID:97077455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1690 DGTEEWRCLLGFKKDDDGNRCVADDAPVCNNNNGGCDKNADCREVENTDRDPSKKIVCTC 1749
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1532 MIELYKPQLNAVNEQI-----AAIGTEP----TDAEKK--KYAPIFEDLKGLYETI 1576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1630 HIASIALNNLNKSGLVTEGESKKILAKMLNMDAMDLLGIGSNHVCISTSTPDNAGCFRYD 1689
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   303 DEREECKCLLNYKQ - - EGDKCEENPNPTCNENNGGCDADATCTEEDS - - GSSRKKITCEC 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      65 LESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEK 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       125 VLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNL 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8 NILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDI---LNSRLKKRKYFLDV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     185 VNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DTKKDMLGKLLSTGLV-QIFPNTIISKLIEGKFQDMLNI-SQHQCVKKQCPENSGCFRHL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1785;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                27.7%; Score 569; DB 2; Lv 33.0%; Pred. No. 1.6e-22; ive 92; Mismatches 138;
                            359 TKPDSYPLFDGIFCSSSNFLGISFLLILMLILYS 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        rhoptry protein - Plasmodium yoelii (fragment)
359 TKPDSYPLFDGIFCSSSNFLGISFLLILMLILYS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 27.73
Best Local Similarity 33.09
Matches 130; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        245
                                                                                                                               RESULT
                                            g
     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             C; Species: Plasmodium yoelli
C; Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 26-Aug-1999
C; Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 26-Aug-1999
C; Accession: A45532; A45531
A; Title: Cloning and analysis of the gene encoding the 230-kilodalton merozoite surface
A; Reference number: A45532; MUID: 90014981
A; Reference number: A45532
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-1772 < LEWA
A; Residues: 1-1772 < LEWA
A; Cross-references: GB: J04668; NID: 9160492; PID: 9160493
A; Title: Precursor to the major merozoite surface antigen of Plasmodium yoelii: cloning
A; Reference number: A45531; MUID: 90014982
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1625 LINILIKSGLVGEGES-----KKILAKMLNMDGMDLLGVDPKHVCVDTRDIPKNAGCF 1676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     : :: | | : | | : | | 475 GQAEEYSEELQNYKNEKAEFEILTKNLEKYIQIDEKLDEFVEHAENNKHIASI--A 532
                                                                                                                                                                                                   300 RHLDEREECKCLLNYKQ-EGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCEC 358
                                                                                                                                                                                                                             DLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKVLA 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKE--SKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    186 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLF----KNTNDFEAIKKL: 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        300 RHLDEREECKCLLNYKQ-EGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCEC 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8 NILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKKKYFLDVLES 67
                                                                                                                               186 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLF----KNTNDFEAIKKL
                                                                                                242 INDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNIS-QHQCV-KKQCPENSGCF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       242 INDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNIS-QHQCV-KKQCPENSGCF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 1772;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                   A45532
major merozoite surface antigen precursor - Plasmodium yoelii
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: DNA
A; Residues: 454-1094 <DAL>
A; Cross-references: GB:J03975; NID:g160081; PID:g160082
C; Superfamily: major merozoite surface antigen
C; Keywords: surface antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28.9%; Score 593.5; DB 2; 32.5%; Pred. No. 8e-24; live 96; Mismatches 137;
                                                                                                                                                                                                                                                                                                                              359 TKPDSYPLFDGIFCSSSNFLGISFLLILMLILYS 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 32.5
Matches 128; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Accession: A45531
A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                   645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    128
                                                                                                                                                                                                                                                     585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 89
                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                             g
                                         g
                                                                                                                                                  g
                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     <u>අ</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
à
                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

11;

34;

for high-molecular ma

```
R; Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Blochem. Parasitol. 42, 241-246, 1990
A; Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple cd
A; Reference number: A45521; MUID:91101660
A; Accession: A45521
A; Mulcoule type: DNA
A; Residues: 2260-2401 < KEE>
A; Cross-references: GB:M34281
                                                                                                                                                                                                                                                                                                    15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   524 SKSLINETKKSIEEEYQNINTLKKVDDYIKVCLNTNEL----ITNCHNKQTTLKDKLNQ 578
                                                                                                                                                                                                                                                                                                                                                                       285 NEIXNTIKSDEDKIYVGDIEQLYNEMFSVVQESNIEHIENKTEILTLKTKIDNVYNNIQN 344
                                                                                                                                                                                                                                                                                                                                                                                                                                  LESDLMQ--FKHISSNEYIIE---DSFKLLNSEQKNIL---LKSYKYIKESVENDI-KFA 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                           116 QEGIS---YYEKVL---AKYKD--DLESIK-KVIKEEKEKFPSSPPTTPPSPAKTDEQKK 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         167 ESKFL--PFLTNIETLYNN------LVANKI-----DDYLINLKAKIND 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             202 ----CNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDDTKKDMLGKLLST 257
                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                           7 DNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDV-- 64
                                                                                                                                                                                                                                                       DB 2; Length 2401;
                                                                                                                                                                                                                                                     Query Match 8.2%; Score 168.5; DB 2; Length 2. Best Local Similarity 22.6%; Pred. No. 0.29; Matches 73; Conservative 68; Mismatches 127; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                    65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
```

8, 2001, 12:30:54

Search completed: August Job time: 102 sec

DP QY •

.

Merozoite surface Merozoite surface PEMSP1(p19)5 prote PEMSP1(p19)5 prote MSP1EGF2A EGR2-lik MSP1EGF2B EGR2-lik MSP1EGF2B EGR2-lik MSP1EGF2B EGR2-lik

P. yoelii combined MSP1EGF1A EGF1-lik MSP1EGF1B EGF1-lik Recombinant vaccin

Plasmodium falcipa Plasmodium falcipa Plasmodium falcipa Plasmodium falcipa Plasmodium falcipa Plasmodium falcipa Plasmodium falcipa

Plasmodium falcipa Plasmodium falcipa Plasmodium falcipa Plasmodium falcipa Plasmodium falcipa

Merozoíte apical-e

Merozite apical-en

Plasmodium falcipa Plasmodium falcipa S. cerevisiae esse Plasmodium falcipa

Plasmodium falcipa Plasmodium falcipa

B. burgdorferi ant

Plasmodium falcipa Plasmodium falcipa burgdorferi ant burgdorferi ant

Title: Perfect score:

Sequence:

OM protein

Run on:

Scoring table:

Searched:

Minimum DB Maximum DB

Database

```
Odink K, Lockyer M, Riveros-Moreno V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence of the P195 protein of Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Malaria vaccine; epitope; antigen; immunogen
                                                                                                                                                                                                                                                                                                                                                                               ALIGNMENTS
                                                                               AAR41354
AAR41355
AAY70278
AAB18253
AAB18218
                                                                                                                                             AAB18291
AAB18294
AAB18236
AAB18275
AAR07503
AAW24575
                                                                                                                                                                                                                   AAB18167
AAB18265
AAB18241
AAB18171
AAY97001
AAB18205
AAB18273
                                                                                                                                                                                                                                                                                            AAB18301
AAY20058
                                                                                                                                                                                                                                                                                                                                                                                                                                  AAP50777 standard; Protein; 1654 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (WELL ) WELLCOME FOUNDATION LTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     84GB-0024340.
84GB-0004692.
85GB-0004429.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    85EP-0301173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sandhu J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1985-224845/37.
N-PSDB; AAN50530.
                                                                                                                                  5.7.7.7.7.7.7.5.5
6.6.6.6.6.8.8.9
5.5.7.7.7.7.7.5.5.5
6.5.6.6.6.6.8.8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      26-SEP-1984;
22-FEB-1984;
21-FEB-1985;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   21-FEB-1985;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30-SEP-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              11-SEP-1985
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EP154454-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Holder A,
 472
472
456.5
302
283
243
243
243
201
198
1186
                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAP50777;
                                                                                                                                                                                                                                                                                                                                                                                                                        AAP50777
 RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                  Merozoite surface
Merozoite surface
Merozoite surface
Modified merozoite
Modified merozoite
Merozoite surface
                                                                   August 8, 2001, 12:31:56; Search time 60.73 Seconds (without alignments) 376.342 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence of the P1
P. falciparum synt
Merozoite surface
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                   AVTISVIDNILSKIENEYEV......SNFLGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                                                                                                                                 /SIDSB/gcgdata/geneseq/geneseqp/AA1982.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1982.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1983.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1985.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1985.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1985.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1981.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1981.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1991.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1992.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1992.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1993.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1995.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1995.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1995.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1995.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1995.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1995.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1995.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1995.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*
                                                                                                                                                                                                                                                                                                                 /SIDS8/gcgdata/geneseq/geneseqp/AA1980.
/SIDS8/gcgdata/geneseq/geneseqp/AA1981.
         4.5
Compugen Ltd
                                                                                                                                                                                                          Total number of hits satisfying chosen parameters:
                                                                                                                                                                                      412676 seqs, 60623988 residues
         GenCore version
Copyright (c) 1993 - 2000
                                                                                                                                                                                                                                                            Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAW54145
AAY09372
AAY05832
AAY05833
AAY05833
AAY09374
AAY05834
AAY05834
AAW36103
AAW36103

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAP50777
                                                                                                                                                       BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                              length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                       A_Geneseq_0601:*
                                                                                                            US-09-500-376-5
1982
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DΒ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0.84444887.60
                                                                                                                                                                                                                                                                                                                                                                                                            Post-processing:
                                                                                                                                                                                                                              sed
```

Cloned DNA sequence encoding plasmodium falciparum protein - useful for expressing the protein for use in vaccines against malaria

PfMSP1(p19)A prote PfMSP1(p19)A prote

1942 1818.5 1726.5 1726.5 1726.5 1726.5 1720.5 1720.5 1720.5 553

10 9 8 7 6 9 8 9 2

Score

Result No.

```
WO9920774-A2
                                                                                                                                                                     Sequence
                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY09372;
                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                             241
                                                                                                                                                                                                                                                                                                                                                                             1504
                                                                                                                                                                                                                                                                                                                                                                                                              1563
                                                                                                                                                                                                                                                                                                                                                                                                                               361
                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                              301
                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                           ó
                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                           QC
                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                           Óλ
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                          ö
                                                                                                                                                                                                                                                    1278 avttsvidnilskieneyevlylkplagvyrslkkglennvmtfnvnvkdilnsrfnkre 1337
                                                                                                                                                                                                                                                                                                            360
                                                                                                                                                                         61 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                          GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                            240
                                                                                                                                                                                                                                                                            300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      gp190; malaria; MSP-1; merozoite surface protein; stability; vaccine;
monoclonal antibody; passive immunisation; parasite.
                                                                                                                           Gaps
                                                                                                                                           9
                     The sequence encoding the P195 protein of Plasmodium falciparum (AANSO530) and a peptide comprising at least one of its epitopes see AARSO777) are claimed. Also claimed is a vaccine for inducing immunity to malaria comprising the novel peptide or P195 or a peptide comprising at least one epitope when derived from the new DNA sequence, together with a carrier.
                                                                                                                                          AVTISVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE
                                                                                                                                                                                                                                                                                                           GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSNF
                                                                                                                                                                                                                                           LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                          VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE
                                                                                                                          ö
                                                                                                          Length 1654;
                                                                                                                          Indels
                                                                                                        Score 1942; DB 6;
Pred. No. 1.7e-130;
0; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       P. falciparum synthetic gp190 protein.
                                                                                                                                                                                                                                                                                                                                                                                                       ¥.
                                                                                                                                                                                                                                                                                                                                                                                                     AAW54145 standard; Protein; 1639
       6; Fig 1; 51pp; English.
                                                                                                        Query Match 98.0%;
Best Local Similarity 98.7%;
Matches 372; Conservative
                                                                                                                                                                                                                                                                                                                                                      1638 lgisfllilmlilysfi 1654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         97WO-EP05441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        96DE-4040817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ä
                                                                                                                                                                                                                                                                                                                                            LGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Tolle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Plasmodium falciparum
                                                                                 1654 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pan W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (BUJA/) BUJARD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO9814583-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         02-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        02-OCT-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                      23-SEP-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        09-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bujard H,
                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                     AAW54145;
       Claim
                                                                                                                                                                                          1338
                                                                                                                                                                                                          121
                                                                                                                                                                                                                          1398
                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                           241
                                                                                                                                                                                                                                                                                                                           1578
                                                                                                                                                                                                                                                                                                                                            361
                                                                                                                                                                                                                                                                                                            301
a
                                                                                                                                                                                        g
                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                       q
                                                                                                                                           õ
                                                                                                                                                                         ò
                                                                                                                                                                                                          ô
                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                           Óγ
                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                          g
```

```
1;
                                                                                                                                                                                                                                                                  (merozoite surface) protein. The gene encoding this protein has been stabilised by reducing the AT content of the nucleotide sequence. Such a protein is useful in vaccines against malaria or for producing monoclonal antibodies (for passive immunisation). The complete gp190 protein can now be produced outside the parasite and has, at least over extended regions, the native pattern of folding. Larger amounts of the protein can be produced recombinantly than would be possible using the parasites as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLITKFLSTGM 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VFENILLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCSSSNF 360
                                                                                                                                                                                                                                               This sequence represents a modified Plasmodium falciparum gp190/MSP-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                          φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 AVTISVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 60
                                                                         Recombinant production of complete gp190/MSP-1 Plasmodium surface protein - useful in anti-malaria vaccines, also stabilising genes reducing their AT content
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 19; Length 1639;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MSP-1; merozoite surface protein; malaria; vaccine; protein engineering; protein expression; codon usage;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. 1.2e-121;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 1818.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Merozoite surface protein MSP-1-42.
                                                                                                                                                                                         Example 1; Fig 3c; 48pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY09372 standard; Protein; 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             91.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               93.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-PSDB; AAV21451, AAV35363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LGISFLLILMLILYSFI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Plasmodium falciparum
1998-240088/21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1639 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              transgenic animal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              31-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Sim
Matches 352;
```

```
Plasmodium falciperum. The nucleir acid sequence encoding MSP-1-42 has been modified (see AAX2586) compered to the native sequence (see AAX25587) such that 306 nucleotide positions have been replaced to lower the AT content (from 76 to 49.7%) and to eliminate 10 mRNA instability motifs while maintaining the same protein amino acid sequence. These alterations allow MSP-1-42 to be expressed in mammalian cell culture and in transgenic mice. The invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   New modified recombinant nucleic acid sequences useful for producing malarial DNA vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                provides modified recombinant nucleic acid sequences and methods for increasing the mRNA levels and protein expression of proteins that are difficult to express in cell culture systems, mammalian cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The present sequence represents a 42 kDa C-terminal portion malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against
                                                                                                 protein engineering; protein expression; codon usage;
                                                                            MSP-1; merozoite surface protein; malaria; vaccine;
                                       Merozoite surface protein MSP-1-42.
                                                                                                                                                                                                                                                                                                                                                                             (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; Fig 1; 43pp; English.
                                                                                                                                                                                                                                                                            98WO-US22225
                                                                                                                                                                                                                                                                                                                   98US-0085649
                                                                                                                                                                                                                                                                                                                                        97US-0062592
02-AUG-1999 (first entry)
                                                                                                                                                          Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-302742/25.
N-PSDB; AAX25586.
                                                                                                                                                                                                                                                                                                                                                                                                                    Chen LH, Meade H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           355 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
                                                                                                                     transgenic animal.
                                                                                                                                                                                               W09920766-A2
                                                                                                                                                                                                                                                                            20-OCT-1998;
                                                                                                                                                                                                                                                                                                                   15-MAY-1998;
                                                                                                                                                                                                                                                                                                                                      20-0CT-1997;
                                                                                                                                                                                                                                       29-APR-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 332;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
                                                                                                                                                                                                                                                                                                                                                                     The present sequence represents a 42 kDa C-terminal portion of malaria merzoite surface protein MSP-1 (i.e. MSP-1-42), an malaria merzoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against plasmodium falciparum. The nucleic acid sequence encoding MSP-1-42 has been modified (see AAX56008) compared to the native sequence (see AAX56009) such that 306 nucleotide positions have been replaced to lower the AT content (from 76 to 49.7%) and to eliminate 10 mRNA instability motifs while maintaining the same protein amino acid sequence. These alterations allow MSP-1-42 to be expressed in mammalian cell culture and in fransgenic mice. Native MSP-1-12 is known to be difficult to express in cell culture systems, mammalian cell culture systems or in transgenic animals. The characteric in the milk of transgenic animals, and also provides a DNA vaccine comprising a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      241 VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 AVTTSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 20; Length 355;
                                                                                                                                                                                                                                                                                            Modified malarial protein for use in anti-malarial vaccines
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 1726.5; DB 20; Lengt
Pred. No. 6.1e-116;
6; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           vector containing the altered MSP-1-42 sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY05832 standard; Protein; 355 AA
                                                                                                                                                        (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                    Example; Fig 1; 35pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      87.1%;
93.3%;
                                                       98WO-US22226
                                                                                          98US-0085649.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best_Local Similarity 93.3
Matches 332; Conservative
                                                                                                                                                                                                                                   WPI; 1999-288313/24
                                                                                                                                                                                                                                                       P-PSDB; AAX56008
                                                       20-0CT-1998;
                                                                                          15-MAY-1998;
20-OCT-1997;
                29-APR-1999
                                                                                                                                                                                               Chen LH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
```

οţ

```
61 nfknvlesdlipykdltssnyvvkdpykflnkekrdkflssynyikdsidtdinfandvl 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 AVTTSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 60
                                                                                                                                              motifs or rare codons relative to the recombinant expression system to be used. The invention allows expression of MSP-1 protein in the milk of transgenic animals, also provides a DNA vaccine comprising a vector containing the altered MSP-1-42 sequence.
culture systems or in transgenic animals. The preferred difficult protein candidates for expression are those derived from lower organisms such as parasites, bacteria and viruses that have DNA coding sequences of high AT content or which have mRNA instability
                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 1726.5; DB 20; Length 355;
Pred. No. 6.1e-116;
6; Mismatches 17; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
9
                                                                                                                                                                                                                                                                                                                                                                                                                                   87.1%;
93.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
```

g ò

AAY05832;

AAY05832

121

셤 ò 염 ò g ŏ QQ ò g ò g

1;

61

121 181 181

g

121

QQ Dp

GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCS 356

301

S Q

LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM

```
300
                                                      This present sequence comprises a 42 kDa C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against Plasmodium falciparum. The C-terminal end of the sequence is modified to include a 6xHis tag. A nucleic acid (see AAX56008) encoding MSP-1-2 has been modified according to a method of the invention in order to improve expression in mammalian cells and in transgenic animals by reducing the AT content and removing mRNA instability motifs. The invention allows expression of MSP-1-42 in the milk of transgenic animals, and also provides a DNA vaccine comprising a vector containing the altered MSP-1-42 nucleic
                                       241 VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE
                                                                                                             301 GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCS 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Modified malarial protein for use in anti-malarial vaccines
                                                                                                                                                                                                                                                                                                                         engineering; protein expression; codon usage;
                                                                                                                                                                                                                                                                                                           MSP-1; merozoite surface protein; malaria; vaccine;
                                                                                                                                                                                                                                                                                 Merozoite surface protein MSP-1-42.
                                                                                                                                                                                             AAY09373 standard; Protein; 361 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Example; Fig 2; 35pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0085649.
                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1999-288313/24.
P-PSDB; AAX56009.
                                                                                                                                                                                                                                                                                                                                        transgenic animal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Meade H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        361 AA;
                                                                                                                                                                                                                                                                                                                                                                                            WO9920774-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                 20-OCT-1998;
                                                                                                                                                                                                                                                    31-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15-MAY-1998;
20-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                      29-APR-1999
                                                                                                                                                                                                                        AAY09373;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Chen LH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                         protein
                                                                                                                                                                                                             음
                                                                                              ŏ
                                                                                                                      Dp
```

```
This present sequence comprises a 42 kDa C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vacche against plasmodium falciparum. The C-terminal end of the sequence is modified to include a 6xHis tag. Nucleic acids (see AAX25586 and AAX25593) encoding MSP-1-42 have been modified according to a method of the invention in order to improve expression in mammalian cells and in transgent animals. The invention provides modified recombinant nucleic acid sequences and methods for increasing the recombinant nucleic acid sequences and methods for increasing the recombinant nucleic acid sequences and methods for increasing the company of protein expression of protein expression are those derived from lower culture systems or in transgenic animals. The preferred difficult protein candidates for expression are those derived from lower conganisms such as parasites, bacteria and viruses that have DNA coding sequences of high AT content or which have mRNA instability motifs or rare codons relative to the recombinant expression system
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               producing
New modified recombinant nucleic acid sequences useful for
                                                                                                                                                                                                                                            MSP-1; merozoite surface protein; malaria; vaccine; protein engineering; protein expression; codon usage;
                                                                                                      Æ
                                                                                                                                                                                                           Merozoite surface protein MSP-1-42.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Disclosure; Fig 2; 43pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                    AAY05833 standard; Protein; 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                    98US-0085649.
                                                                                                                                                                                                                                                                                                                                                                                                                     98WO-US22225
                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 nalarial DNA vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-302742/25.
                                                                                                                                                                                                                                                                                transgenic animal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Chen LH, Meade H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-PSDB; AAX25587
                                                                                                                                                                                                                                                                                                                                                                                                                     20-OCT-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      15-MAY-1998;
20-OCT-1997;
                                                                                                                                                                         02-AUG-1999
                                                                                                                                                                                                                                                                                                                                                 WO9920766-A2
                                                                                                                                                                                                                                                                                                                                                                                   29-APR-1999
                                                                                                                                       AAY05833;
                                                                    9
                  300
```

1;

Gaps

87.1%; Score 1726.5; DB 20; Length 361; llarity 93.3%; Pred. No. 6.3e-116; Conservative 6; Mismatches 17; Indels 1;

Best Local Similarity Matches 332; Conserv

Query Match

NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120

61

ò

S G

S

```
Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY05834;
                   ĽH,
                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         œ
                    Chen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
AAY05834
ID AAY0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                 ij
                                                                                                                                                                                                                                                                                             GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                               LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VFENILKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE 300
                                                                                                                                                                                                                                                                          NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                 Gaps
                                                                                                                                                                                                 1 AVTISVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 60
                                                                                                                                                                                                                    to be used. The invention allows expression of MSP-1 in the milk of transgenic animals, and also provides a DNA vaccine comprising vector containing the altered MSP-1-42 nucleic acid.
                                                                                                                                                                                                                                                                                                                                                                                                                                        /note= "Asn·in native MSP-1-42 (N262Q mutation)"
                                                                                                                                                                 1;
                                                                                                                           DB 20; Length 361;
                                                                                                                       Score 1726.5; DB 20; Lengt
Pred. No. 6.3e-116;
6; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   'note- "Asn in native MSP-1-42 (N1810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MSP-1; merozoite surface protein; malaria; vaccine; protein engineering; protein expression; codon usage; transgenic animal; mutant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /note= "beta-casein signal peptide'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Modified merozoite surface protein MSP-1-42.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            371..376
/note= "6xHis tag"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY09374 standard; Protein; 376
                                                                                                                           87.1%;
93.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98US-0085649.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97US-0062592.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                         Query Match 87.1
Best Local Similarity 93.3
Matches 332; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Plasmodium falciparum
                                                                        361 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              W09920774-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       31-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20-0CT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29-APR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY09374;
                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pept1de
                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
AAY09374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HANDER STANDER STANDER
SSSSXS
                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
```

```
NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ţ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 AVTTSVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE 60
                                                                                                                                                                                                                    The present sequence represents a 42 kDa C-terminal portion of manaria merozoite surface protein MSP-1 (1.e. MSP-1-42), an important target for the development of a vaccine against Plasmodium falciparum. The sequence has been modified to include an N-terminal beta-casein signal peptide and an N-terminal oxHis and N262, mutations have been introduced eliminate 2 N-01ycosylation sites. These modifications allow the MSP-1-42 protein to be expressed in the milk of transgenic mice.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  301 GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCS 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 376;
                                                                                                                        Modified malarial protein for use in anti-malarial vaccines
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein engineering; protein expression; codon usage;
transgenic animal; mutant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ______beta-casein signal peptide"
371..376
/note= "6xHis tag"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MSP-1; merozoite surface protein; malaria; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. 1.8e-115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Modified merozoite surface protein MSP-1-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           86.8%; Score 1720.5;
93.0%; Pred. No. 1.8e
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
1..15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY05834 standard; Protein; 376
                                                                                                                                                                        Example; Fig 11; 35pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches 331; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Plasmodium falciparum
                                                 WPI; 1999-288313/24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                         376 AA;
Meade H;
                                                                         P-PSDB; AAX56008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         02-AUG-1999
```

;;

315 gdkcvenpnptcnennggcdadakcteedsgsngkkitcectkpdsyplfdgifcs 370

g

ဖ

```
The present sequence represents a 42 kDa C-terminal portion of malaria mercozite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against plasmodium falciparum. The sequence has been modified to include plasmodium falciparum. The sequence has been modified to include an N-terminal beta-casein signal peptide and an N-terminal 6xHis tag. In addition, N1810 and N2620 mutations have been introduced to eliminate 2 N-glycosylation sites. These modifications allow the MSP-1-42 protein to be expressed in the milk of transgenic mice. The invention generally relates to modified recombinant nucleic acid sequences and methods for increasing the mRNA levels and cell culture systems, ammanian cell culture systems or in transgenic animals. The preferred difficult protein candidates for transgenic animals. The preferred difficult protein candidates for transgenic animals. The preferred difficult protein candidates conspicult and viruses that have DNA coding sequences of high AT content or which have mRNA instability motifs or rare codons relative to the recombinant expression system to be used.
                                                                                                                                                                                                                                                                                                                                                                      recombinant nucleic acid sequences useful for producing vaccine
                                                   /note= "Asn in native MSP-1-42 (N2620 mutation)"
                 note= "Asn in native MSP-1-42 (N181Q mutation)"
                                                                                                                                                                                                                                                           (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                               Example; Fig 11; 43pp; English.
                                                                                                                                                                  98WO-US22225
                                                                                                                                                                                                    98US-0085649.
                                                                                                                                                                                                                                                                                                                                  WPI; 1999-302742/25.
Misc-difference 197
                                                                                                                                                                                                                                                                                                  Meade H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                376 AA;
                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAX25593
                                    Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                      New modified
malarial DNA
                                                                                         WO9920766-A2
                                                                                                                                                                20-OCT-1998;
                                                                                                                                                                                                    15-MAY-1998;
20-OCT-1997;
                                                                                                                             29-APR-1999
                                                                                                                                                                                                                                                                                                Chen LH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
```

; 255 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180 LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 240 VFENLLKSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQE 300 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120 Gaps AVTISVIDNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSRFNKRE '60 GDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCS 356 1; DB 20; Length 376; Score 1720.5; DB 20; Lengt. Pred. No. 1.8e-115; 5; Mismatches 19; Indels Query Match 86.8%; Best Local Similarity 93.0%; Matches 331; Conservative 9/ 121 136 181 196 241 256 301 ò g g g õ a ò Q à ò

ö

240

LNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM

181

Ω g

121 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180 

NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120

61

ò

g à QQ

```
fragment of a Plasmodium merosotic surface protein-1 (MSP-1). The non-natural variants have reduced affinity for at least 1 antibody capable of blocking a second antibody that inhibits the proteolytic cleavage of blasmodium MSP-1_4_2, and has the same affinity for at least one third antibody that inhibits the proteolytic cleavage of plasmodium MSP-1_4_2, and proteolytic cleavage of plasmodium present invention are useful for immunising a mammal against malaria, and can be used to treat malaria. The present sequence is MSP-133 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The present invention relates to non-natural variants of a C-terminal
                                                                                                                                                                                                                                                                                                                                                                     Uthaipibull C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 AVTTSVIDNILSKIENEYEVLYLKPLAGVYRSLKKOLENNVMTFNVNVKDILNSRFNKRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Novel variants of the C-terminal fragment of Plasmodium merozoite surface protein-1, useful as vaccines for treating or preventing malaria -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 262;
                                                                                                                                           Merozoite surface protein; protazoacide; vaccine; malaria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                      ŝ
                                                                                                                                                                                                                                                                                                                                                                      Syed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         64.7%; Score 1282; DB 22;
97.7%; Pred. No. 2.6e-84;
live 1; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                      Morgan W,
                             Ä.
                                                                                                                                                                                                                                                                                                                                                                      ה,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Example 5; Fig 15; 126pp; English.
                          AAB37610 standard; Protein; 262
                                                                                                                                                                                                                                                                                                                                                                      Feeney
                                                                                                               surface protein-133
                                                                                                                                                                                                                                                                                 99GB-0009072.
99US-0311817.
99CA-2271451.
                                                                                                                                                                                                                                                                                                                                          (MEDI-) MEDICAL RES COUNCIL.
                                                                                                                                                                                                                                                        20-APR-2000; 2000WO-GB01558
                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity 97.7 ses 252; Conservative
                                                                                                                                                                                                                                                                                                                                                                     Holder A, Birdsall B,
                                                                                                                                                                    Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2001-015762/02.
N-PSDB; AAC68978.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   262 AA;
                                                                                                                                                                                                 WO200063245-A2.
                                                                                                                                                                                                                                                                                  20-APR-1999;
13-MAY-1999;
                                                                                                                                                                                                                                                                                                              25-MAY-1999;
                                                                                   27-FEB-2001
                                                                                                                                                                                                                            26-OCT-2000
                                                       AAB37610;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Loca
Matches
RESULT
AAB37610
                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
```

```
/note= "glycosylphosphatidylinositol anchoring sequence"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This is the amino acid sequence of a recombinant protein comprising amino acids 1613-1705 of the Plasmodium falciparum merozoite surface protein 1 (MSP1) 19 kD C-terminal fragment (p19), linked to a glycosylphosphatidylinositol membrane anchoring sequence. p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the p19 fragment provides a high level of protective immunity since it includes epitopes not presented in the p42 fragment.
                                                                                                                                                                                                                                                                                                                                                                                /note= "amino acids derived from P. falciparum MSP1 p19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           272 KTMIQQSSGCFRHLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSG 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Recombinant protein containing Plasmodium merozoite surface protein-1 p42 fragment – useful in antimalarial vaccines, also new antibodies for diagnosis and protein purification
                                                                                                                                                                                                                                      Plasmodium vivax; merozoite surface protein; MSP1; p19;
Plasmodium falciparum; malaria; vaccine; immunity; epitope.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 71 sngkkitcectkpdsyplfdgifcsssnflgisfllilmlilysfi 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 553; DB 18;
Pred. No. 1.3e-32;
2; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Barnwell JW, Longacre-Andre S, Mendis K,
                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers
                                                                                        AAW22592 standard; Protein; 116 AA
                                                                                                                                                                                                                                                                                                                                                                                                      fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Disclosure; Fig 1B; 85pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5;
                                                                                                                                                                                                  PfMSP1(p19)A protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     27.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  97WO-FR00291.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      96FR-0001821.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (UYNY ) UNIV NEW YORK STATE.
                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                         Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PASTEUR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1997-425034/39.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
es 99; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    P-PSDB; AAW22592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INSP ) INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO9730159-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      14-FEB-1996;
                                                                                                                                                              25-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21-AUG-1997
                                                                                                                                                                                                                                                                                                               Synthetic.
                                                                                                                            AAW22592;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                    Region
                                                                                                                                                                                                                                                                                                                                                                                                                       Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Roth
                                                                                                                                                                                                                                                                                                                                                 Key
                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note= "glycosylphosphatidylinositol anchoring sequence"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This is the amino acid sequence of a recombinant protein comprising amino acids 1613-1705 of the Plasmodium falciparum merozoite surface protein 1 (MSP1) 19 kD C-terminal fragment (p19), linked to a glycosylphosphatidylinositol membrane anchoring sequence. p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the p19 fragment provides a high level of protective immunity since it includes epitopes not presented in the p42 fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                  /note= "amino acids derived from P. falciparum MSP1 p19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             272 KTMIQQSSGCFRHLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSG 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        nt protein containing the merozoite surface protein-1 p19 - useful in anti-malarial vaccines, diagnosis and protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 553; DB 18; Length 116;
Pred. No. 1.3e-32;
2; Mismatches 5; Indels (
                                                                                                                                                                                                                                                                      Plasmodium vivax; merozoite surface protein; MSP1; p19;
Plasmodium falciparum; malaria; vaccine; immunity; epitope.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 <u>г</u>,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Nato
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mendis K,
                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers
                                                                                                                      AAW36103 standard; Protein; 116 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                        fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Disclosure; Fig 1B; 85pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Longacre-Andre S,
                                                                                                                                                                                                                                    PfMSP1(p19)A protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 27.9%;
Best Local Similarity 93.4%;
Matches 99; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  97WO-FR00290.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (INSP ) INST PASTEUR.
(UYNY ) UNIV NEW YORK STATE.
                  241 vfenlaktvlsnlldgnl 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      96FR-0001822
241 VFENLLKSVLSNLLDWKL 258
                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                             Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1997-425033/39.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Recombinant protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  116 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAT94550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Barnwell JW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          purification
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO9730158-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      14-FEB-1996;
                                                                                                                                                                                                25-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21-AUG-1997.
```

Roth C;

Synthetic

Region

Region

AAW36103;

g

ð

Nato F;

ö

Gaps

ö

Indels

g

SNGKKITCECTKPDCYPLFDGIFCSSSNFLGISFLLILMLILYSFI 377

332

ð 염 ò

Seguence

Length 116;

Uthaipibull C;

ŝ Syed

AAB37608

12

AAB37608 RESULT

```
The present invention relates to non-natural variants of a C-terminal fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The mon-natural variants have reduced affinity for at least 1 antibody capable of blocking a second antibody that inhibits the proteclytic cleavage of Plasmodium MSP-1_4_2, and has the same affinity for at least one third antibody that inhibits the proteclytic cleavage of Plasmodium MSP-1_4_2. The non-natural variants of the present invention are useful for immunishing a mammal against malaria, and can be used to treat malaria. The present sequence is MSP-119 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               272 KTMIQQSSGCFRHLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSG 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Novel variants of the C-terminal fragment of Plasmodium merozoite surface protein-1, useful as vaccines for treating or preventing malaria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Chimeric; Plasmodium vivax; merozoite surface protein; MSP1; p19;
Plasmodium falciparum; malaria; vaccine; immunity; epitope.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     23.8%; Score 472; DB 22; Length 108; 93.2%; Pred. No. 7.6e-27; tive 1; Mismatches 5; Indels
                                    Merozoite surface protein; protazoacide; vaccine; malaria.
                                                                                                                                                                                                                                                                                                                                          Morgan W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          332 SNGKKITCECTKPDCYPLFDGIFCSSSN 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
1..19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAW22593 standard; Protein; 127 AA.
                                                                                                                                                                                                                                                                                                                                          Feeney J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Example 5; Fig 15; 126pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              - Plasmodium vivax.
- Plasmodium falciparum.
Merozoite surface protein-119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PfMSP1(p19)S protein sequence.
                                                                                                                                                                                                                          99GB-0009072.
                                                                                                                                                                                       20-APR-2000; 2000WO-GB01558.
                                                                                                                                                                                                                                                               99CA-2271451
                                                                                                                                                                                                                                                                                                      (MEDI-) MEDICAL RES COUNCIL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 93.2
Matches 82; Conservative
                                                                                                                                                                                                                                                                                                                                       Holder A, Birdsall B,
                                                                         Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                            WPI; 2001-015762/02.
N-PSDB; AAC68977.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ¥,
                                                                                                             WO200063245-A2
                                                                                                                                                                                                                            20-APR-1999;
                                                                                                                                                                                                                                                               25-MAY-1999;
                                                                                                                                                                                                                                               13-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25-MAR-1998
                                                                                                                                                  26-OCT-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Chimeric
Chimeric
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAW22593;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAW22593
   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The non-natural variants have reduced affinity for at least 1 antibody capable of blocking a second antibody that inhibits the proteolytic cleavage of Plasmodium MSP-1_4_2, and has the same affinity for at least one third antibody that inhibits the proteolytic cleavage of Plasmodium MSP-1_4_2, and has the same affinity for at least one third antibody that inhibits the proteolytic cleavage of Plasmodium MSP-1_4_2. compared to natural MSP-1_1_9. The present sequence is the wild-type MSP-1 protein. This sequence was used to generate the variants of the present invention. The non-natural variants of the present invention. The mon-natural variants of the present invention are useful for immunising a mammal against malaria, and can be
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ပ်
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    272 KTMIQQSSGCFRHLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSG 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Uthaipibull
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Novel variants of the C-terminal fragment of Plasmodium merozoite surface protein-1, useful as vaccines for treating or preventing malaria -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 96;
                                                                                                                                                                                                        surface protein; protazoacide; vaccine; malaria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 472; DB 22; Length 9
Pred. No. 6.5e-27;
1; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Syed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Morgan
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          332 SNGKKITCECTKPDCYPLFDGIFCSSSN 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Example 2; Page 48; 126pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB37609 standard; Protein; 108 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ر.
د
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sngkkitcectkpdsyplfdgifcssn
                                                  AAB37608 standard; protein; 96 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Feeney
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23.8%;
93.2%;
                                                                                                                                                                                                                                                                                                                                                                                           99GB-0009072.
                                                                                                                                                                                                                                                                                                                                                                                                             99US-0311817.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (MEDI-) MEDICAL RES COUNCIL.
                                                                                                                                                                 Merozoite surface protein-1.
                                                                                                                                                                                                                                                                                                                                                      20-APR-2000; 2000WO-GB01558
                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 Similarity 93.2
82; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Birdsall B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                used to treat malaria
                                                                                                                                                                                                                                           Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2001-015762/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                WO200063245-A2.
                                                                                                                              27-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                           20-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                               .3-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                              25-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     27-FEB-2001
                                                                                                                                                                                                                                                                                                                   26-OCT-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Holder A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
```

Query Match

Best Local

Matches

qq

ŏ Q

ò

AAB37609

axaxax xaxaxa

AAB37609

69

ö

Gaps

.; O

σ

```
- Plasmodium vivax.
- Plasmodium falciparum.
  Chimeric
Chimeric
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                            Peptide
                                                                                                                  Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Roth C;
                                                                                                                                                        Region
                                                                                                                                                                                               Region
                                                                                                                                                                                                                                                         Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 78
                                                              Key
    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This is the amino acid sequence of a chimeric protein comprising amino acids 1-32 of the Plasmodium vivax merozoite surface protein 1 (MSP1) linked to amino acids Asn1613-Ser1705 which correspond to the 19 KD C-terminal fragment (P19) of MSP-1 from Plasmodium falciparum. P19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the P19 fragment provides a high level of protective immunity since it includes epitopes not presented in the P42 fragment.
                                                                                                                    З
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  247 KSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQEGDKCVE 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 :::|: :: :::: critical control of the control of 
                                                                                                                "encoded by restriction enzyme sequence used create the chimeric sequence"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Recombinant protein containing Plasmodium merozoite surface protein-1 p42 fragment - useful in antimalarial vaccines, also new antibodies for diagnosis and protein purification
                                                                                                                                                                     /note= "derived from P. falciparum C-terminal p19 fragment of MSP1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Chimeric; Plasmodium vivax; merozoite surface protein; MSP1; p19; Plasmodium falciparum; malaria; vaccine; immunity; epitope.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 127;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCS 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                          /note= "derived from P. vivax MSP1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Nato F;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     23.0%; Score 456.5; DB 18; 72.7%; Pred. No. 1.2e-25; ive 12; Mismatches 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mendis K,
/note= "signal peptide"
                                      "mature protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAW36102 standard; Protein; 127 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Fig 1C; 85pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Longacre-Andre S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PfMSP1(p19)S protein sequence
                                                                                                                                                                                                                                                                                                         97WO-FR00291
                                                                                                                                                                                                                                                                                                                                                  96FR-0001821
                                                                                                                                                                                                                                                                                                                                                                                                            NEW YORK STATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 23.0
Best Local Similarity 72.7
Matches 80; Conservative
                  20..127
                                                                                                                                                    35..127
                                                                                            33..34
/note=
                                        /note=
                                                                                                                                                                                                                                                                                                                                                                                          PASTEUR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1997-425034/39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             127 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               P-PSDB; AAW22592
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Barnwell JW,
                                                                                                                                                                                                                                                                                                                                                                                        (INSP ) INST
(UYNY ) UNIV
                                                                                                                                                                                                                                 WO9730159-A2
                                                                                                                                                                                                                                                                                                           14-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                  14-FEB-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25-MAR-1998
                                                                                                                                                                                                                                                                       21-AUG-1997.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAW36102;
                      Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Roth C;
                                                        Region
                                                                                                                                                      Region
                                                                                              Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAW36102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

```
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     acids 1-32 of the Plasmodium vivax merozoite surface protein 1 (MSP1) linked to amino acids Asn1613-Ser1705 which correspond to the 19 kD c-terminal fragment (p19) of MSP-1 from Plasmodium falciparum. p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the p19 fragment provides a high level of protective immunity since it includes epitopes not presented in the p42 fragment.
                                                                                                                                                                                                       /note= "encoded by restriction enzyme sequence used to
   create the chimeric sequence"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This is the amino acid sequence of a chimeric protein comprising amino
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          247 KSVLSNLLDWKLARYVKHFTTPMRKKTMIQQSSGCFRHLDEREECKCLLNYKQEGDKCVE 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Recombinant protein containing the merozoite surface protein-1 p19 fragment - useful in anti-malarial vaccines, diagnosis and protein purification
                                                                                                                                                                                                                                                                              /note= "derived from P. falciparum C-terminal p19
fragment of MSP1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 127;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTKPDCYPLFDGIFCS 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note= "derived from P: vivax MSP1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nato F;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 456.5; DB 1
Pred. No. 1.2e-25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Longacre-Andre S, Mendis K,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12; Mismatches
                                                                                                   'note= "mature protein"
                                                 'note= "signal peptide"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Search completed: August 8, 2001, 12:31:57 Job time: 165 sec
Location/Qualifiers
1..19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Fig 1C; 85pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         23.0%;
72.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (INSP ) INST PASTEUR.
(UYNY ) UNIV NEW YORK STATE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              97WO-FR00290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                96FR-0001822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1997-425033/39.
N-PSDB; AAT94549.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  127 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Barnwell JW,
                                                                                                                                                                                                                                                                                                                                                          WO9730158-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                            14-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                14-FEB-1996;
                                                                                                                                                                                                                                                                                                                                                                                                          21-AUG-1997.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           80:
```

THIS PAGE BLANK (USPTO)

Merozoite surface P. yoelii combined MSPIEGFIA EGFI-lik MSPIEGFIB EGFI-lik Recombinant vaccin

PfMSP1(p19)S PfMSP1(p19)S MSP1EGF2A E Plasmodium falcipa Plasmodium falcipa

B. burgdorferi ant Plasmodium falcipa Plasmodium falcipa Plasmodium falcipa

Plasmodium falcipa Merozite apical-en Merozoite apical-e

S. cerevisiae esse Plasmodium falcipa Plasmodium falcipa

Plasmodium falcipa Plasmodium falcipa Plasmodium falcipa Plasmodium falcipa

of

Total number Searched:

Minimum DB seq Maximum DB seq

Database

Title: Perfect score: Sequence:

<u>e</u>

Scoring table:

```
DNA sequence encoding plasmodium falciparum protein \mbox{-} for expressing the protein for use in vaccines against
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Odink K, Lockyer M, Riveros-Moreno V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   the P195 protein of Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Malaria vaccine; epitope; antigen; immunogen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ALIGNMENTS
                                       AAW36102
AAR41356
AAR41356
AAR41358
AAR41358
AAR41354
AAR1354
AAR1357
AAR18275
AAR18275
AAR18275
AAR18277
AAR18273
AAR18273
AAR18273
AAR18273
AAR18272
AAR18272
AAR18272
AAR18272
AAR18272
AAR18272
AAR18272
AAR18272
AAR18273
                                                                                                                                                                                                                                                                                                                                                                                                                           AAB18172
AAR07503
AAW24575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAP50777 standard; Protein; 1654 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (WELL ) WELLCOME FOUNDATION LTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               84GB-0024340.
84GB-0004692.
85GB-0004429.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       85EP-0301173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                            Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Holder A, Sandhu J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1985-224845/37
N-PSDB; AAN50530.
26-SEP-1984;
22-FEB-1984;
21-FEB-1985;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21-FEB-1985;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        30-SEP-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11-SEP-1985
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EP154454-A.
 447
447
432.5
432.5
278
278
232
232
232
232
197
197
1158
1158
                                                                                                                                                                                                        142.5
142.5
141.5
139.5
                                                                                                                                                                                                                                                                          137
136.5
136.5
136.5
132
132
131
130.5
129.5
129.5
129.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAP50777;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Cloned DN
useful fo
malaria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAP50777
  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    P. falciparum synt
Merozoite surface
Merozoite surface
Merozoite surface
Merozoite surface
Modified merozoite
Merozoite surface
                                                                                                            (without alignments)
374.345 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PfMSP1(p19)A prote
PfMSP1(p19)A prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence of the P1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                             US-09-500-376-4
1967
1 AVTPSVIHNILSKIENEYEV......SNFLGISFLLILMLILYSFI 375
                                                                                            ; Search time 60.73 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Description
                                                                                                                                                                                                                                                                                                                                                                                                         2: /SIDSB/gcgdata/geneseq/geneseqp/AA1980.DAT:*
3: /SIDSB/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDSB/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDSB/gcgdata/geneseq/geneseqp/AA1982.DAT:*
3: /SIDSB/gcgdata/geneseq/geneseqp/AA1983.DAT:*
3: /SIDSB/gcgdata/geneseq/geneseqp/AA1985.DAT:*
3: /SIDSB/gcgdata/geneseq/geneseqp/AA1986.DAT:*
3: /SIDSB/gcgdata/geneseq/geneseqp/AA1980.DAT:*
3: /SIDSB/gcgdata/geneseq/geneseqp/AA1980.DAT:*
3: /SIDSB/gcgdata/geneseq/geneseqp/AA1981.DAT:*
4: /SIDSB/gcgdata/geneseq/geneseqp/AA1991.DAT:*
4: /SIDSB/gcgdata/geneseq/geneseqp/AA1991.DAT:*
4: /SIDSB/gcgdata/geneseq/geneseqp/AA1991.DAT:*
4: /SIDSB/gcgdata/geneseq/geneseqp/AA1991.DAT:*
4: /SIDSB/gcgdata/geneseq/geneseqp/AA1991.DAT:*
4: /SIDSB/gcgdata/geneseq/geneseqp/AA1995.DAT:*
4: /SIDSB/gcgdata/geneseq/geneseqp/AA1995.DAT:*
4: /SIDSB/gcgdata/geneseq/geneseqp/AA1995.DAT:*
4: /SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
4: /SIDSB/gcgdata/geneseq/geneseqp/AB1999.DAT:*
4: /SIDSB/gcgdata/geneseq/geneseqp/AB1999.DAT:*
4: /SIDSB/gcgdata/geneseq/geneseqp/AB1999.DAT:*
              GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                       hits satisfying chosen parameters:
                                                                                                                                                                                                                                              412676 seqs, 60623988 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                                                                                                                                                                                                                                                                       Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAW54145
AAY09372
AAY05832
AAY05833
AAY05833
AAY05834
AAY05834
AAX05374
AAX050374
AAW36103
                                                                 protein - protein search, using sw mode]
                                                                                            August 8, 2001, 12:31:55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAP50777
                                                                                                                                                                                                      BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                    length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                  A_Geneseq_0601:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1654
1639
355
355
361
361
376
376
262
116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               84.7
84.7
84.7
84.4
84.4
63.3
26.8
```

Score

. 9 Result

1757.5 1665.5 1665.5 1665.5 1665.5 1659.5 1246 528

```
N-PSDB; AAV21451, AAV35363.
         WPI; 1998-240088/21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO9920774-A2
                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY09372;
                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                          m
                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY09372
                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                    Qγ
                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                              ·δ
                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                       ï
                                                                                                                                                                                                                                             1278 avttsvidnilskieneyevlylkplagvyrslkkglennvmtfnvnvkdilnsrfnkre 1337
                                                                                                                                                                              NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                     GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                    298
                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                          gp190; malaria; MSP-1; merozoite surface protein; stability; vaccine; monoclonal antibody; passive immunisation; parasite.
                      The sequence encoding the P195 protein of Plasmodium falciparum (AAN50530) and a peptide comprising at least one of its epitopes (see AAP50777) are claimed. Also claimed is a vaccine for inducing immunity to malaria comprising the novel peptide or P195 or a peptide comprising at least one epitope when derived from the new DNA sequence, together with a carrier.
                                                                                                                                      AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE
                                                                                                                                                                                                                                     LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                    VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE
                                                                                                                       5;
                                                                                                      Length 1654;
                                                                                                                       Indels
                                                                                                     Score 1901; DB 6;
Pred. No. 1.6e-135;
3; Mismatches 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                         falciparum synthetic gp190 protein
                                                                                                                                                                                                                                                                                                                                                                                           Ą.
                                                                                                                                                                                                                                                                                                                                                                                          AAW54145 standard; Protein; 1639
        Claim 6; Fig 1; 51pp; English
                                                                                                       96.6%;
97.1%;
                                                                                                                                                                                                                                                                                                                                            97WO-EP05441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         96DE-4040817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          χ;
                                                                                                                                                                                                                                                                                                                                   LGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Tolle
                                                                                                              Local Similarity 97.1 nes 366; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pan W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (BUJA/) BUJARD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO9814583-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         02-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         02-OCT-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                          23-SEP-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ή,
                                                                                                                                                                                                                                                                                                                                                                                                         AAW54145;
                                                                                Sequence
                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bujard
                                                                                                                                                                                                                   1398
                                                                                                                      Matches
                                                                                                                                                                                     1338
                                                                                                                                                                                                    121
                                                                                                                                                                                                                                                                                                                                   359
                                                                                                                                                                      61
                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                    239
                                                                                                                                                                                                                                                                                                                                                   1638
                                                                                                                                                                                                                                                                                                                                                                                  AAW54145
ð
                                                                                                                                                     g
                                                                                                                                                                    οy
                                                                                                                                                                                 g
                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                 요
                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                   ò
```

```
This sequence represents a modified Plasmodium falciparum gp190/MSP-1 (merozolite surface) protein. The gene encoding this protein has been stabilised by reducing the AT content of the nucleotide sequence. Such a protein is useful in vaccines against malaria or for producing monoclonal antibodies (for passive immunisation). The complete gp190 protein can now the produced outside the parasite amounts of the protein can be produced setting than would be possible using the parasites as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                         ρχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       239 VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 19; Length 1639;
Recombinant production of complete gp190/MSP-1 Plasmodium su
protein – useful in anti-malaria vaccines, also stabilising
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 1757.5; DB 19; Lengt
Pred. No. 1.2e-124;
7; Mismatches 23; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MSP-1; merozoite surface protein; malaria; vaccine; protein engineering; protein expression; codon usage;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Merozoite surface protein MSP-1-42.
                                                                                                    Example 1; Fig 3c; 48pp; German
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY09372 standard; Protein; 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     89.3%;
ilarity 91.2%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             359 LGISFLLILMLILYSFI 375
                                                reducing their AT content
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                     1639 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
Matches 344; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       transgenic animal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31-AUG-1999
```

ñ

```
New modified recombinant nucleic acid sequences useful for producing
                                                                                protein engineering; protein expression; codon usage;
                                                               MSP-1; merozoite surface protein; malaria; vaccine;
                                 Merozoite surface protein MSP-1-42.
                                                                                                                                                                                                                                                                                                         (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; Fig 1; 43pp; English.
                                                                                                                                                                                                                           98WO-US22225.
                                                                                                                                                                                                                                                           98US-0085649.
02-AUG-1999 (first entry)
                                                                                                                             Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                          malarial DNA vaccine
                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-302742/25.
                                                                                               transgenic animal.
                                                                                                                                                                                                                                                                                                                                           Meade H;
                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAX25586
                                                                                                                                                             WO9920766-A2.
                                                                                                                                                                                                                           20-OCT-1998;
                                                                                                                                                                                                                                                                           20-OCT-1997;
                                                                                                                                                                                                                                                           15-MAY-1998;
                                                                                                                                                                                             29-APR-1999
                                                                                                                                                                                                                                                                                                                                           Chen LH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ,;
;;
                                                                                                                                                                                                                                                                                                 The present sequence represents a 42 kba C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against Plasmodium falciparum. The nucleic acid sequence encoding MSP-1-42 has been modified (see AAX56008) compared to the native sequence (see AAX56009) such that 306 nucleotide positions have been replaced to lower the AT content (from 76 to 49.7%) and to eliminate 10 mRNA instability motifs while maintaining the same protein amino acid sequence. These alterations allow MSP-1-42 to be expressed in mammalian cell culture and in transgenic mice. Native MSP-1-12 is known to be difficult to express in cell culture systems.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               mammalian cell culture systems or in transgenic animals. The invention allows expression of MSP-1 protein in the milk of transgenic animals, and also provides a DNA vaccine comprising a vector containing the altered MSP-1-42 sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 1665.5; DB 20; Length 355;
Pred. No. 1.5e-118;
7; Mismatches 22; Indels 3;
                                                                                                                                                                                                                                        Modified malarial protein for use in anti-malarial vaccines
                                                                                                                           (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAX05832 standard; Protein; 355
                                                                                                                                                                                                                                                                          Example; Fig 1; 35pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 91.0%;
Matches 324; Conservative
                                                                            98US-0085649.
97US-0062592.
                                             98WO-US22226
                                                                                                                                                                                         WPI; 1999-288313/24.
P-PSDB; AAX56008.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
                                                                                                                                                           Chen LH, Meade H;
                                             20-OCT-1998;
                                                                            15-MAY-1998;
20-OCT-1997;
               29-APR-1999
```

Sequence

61

g

ò 셤 ò 181

ò

g ò a ð

g

239

```
,;
;
The present sequence represents a 42 kba C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against propertial target for the development of a vaccine against propertial target for the development of a vaccine against the placed to plasmodium falciparum. The nucleic acid sequence encoding MSP-1-42 has been modified (see AAX5586) compared to the native sequence (see AAX5587) such that 306 nucleotide positions have been replaced to lower the AT content (from for 0.7%) and to eliminate 10 mRNA instability motifs while maintaining the same protein amino acid sequence. These alterations allow MSP-1-42 to be expressed in mammalian cell culture and in transgenic mice. The invention provides modified recombinant nucleic acid sequences and methods for increasing the mRNA levels and protein expression of proteins that increasing the mRNA levels and protein expression of protein cell culture systems or in transgenic animals. The preferred difficult protein candidates for expression are those derived from lower coding sequences of high AT content or which have mRNA instability motifs or rare codons relative to the recombinant expression system to be used. The invention allows expression of MSP-1 protein in the milk of transgenic animals, and also provides a DNA vaccine comprising a vector containing the altered MSP-1-42 sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ouery Match 84.7%; Score 1665.5; DB 20; Length 355; Best Local Similarity 91.0%; Pred. No. 1.5e-118; Matches 324; Conservative 7; Mismatches 22; Indels 3;
```

ŏ

AAY05832,

AAY05832
ID AAY0
XX
AC AAY0
XX

300

299

LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM

```
239 VFENLLKSILSNILDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE 298
                                    121
                                                                                  181
                                                         g
                                                                                                     QQ
                                                                                                                                                        엄
                                                                                                                                                                               ò
                                                                                                                                                                                                       q
                                    ò
                                                                                                                                  ò
This present sequence comprises a 42 kDa C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against blasmodium falciparum. The C-terminal end of the sequence is modified to include a 6xHis tag. A nucleic acid (see AAX56008) encoding MSP-1-2 has been modified according to a method of the invention in order to improve expression in mammalian cells and in transgenic animals by reducing the AT content and removing mRNA instability motifs. The invention allows expression of MSP-1-42 in the milk of transgenic animals, and also provides a DNA vaccine comprising a vector containing the altered MSP-1-42 nucleic
                                   239 VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE
                                                                                               299 GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCS 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    protein for use in anti-malarial vaccines
                                                                                                                                                                                                                                                                             engineering; protein expression; codon usage;
                                                                                                                                                                                                                                                                  MSP-1; merozoite surface protein; malaria; vaccine;
                                                                                                                                                                    AAY09373 standard; Protein; 361 AA
                                                                                                                                                                                                                                          Merozoite surface protein MSP-1-42.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Example; Fig 2; 35pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                             98US-0085649.
                                                                                                                                                                                                                                                                                                                                                                                        98WO-US22226
                                                                                                                                                                                                                   (first.entry)
                                                                                                                                                                                                                                                                                                               Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1999-288313/24
                                                                                                                                                                                                                                                                                           transgenic animal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Chen LH, Meade H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Modified malarial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            361 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P-PSDB; AAX56009
                                                                                                                                                                                                                  31-AUG-1999
                                                                                                                                                                                                                                                                                                                                       409920774-A2
                                                                                                                                                                                                                                                                                                                                                                                      20-OCT-1998;
                                                                                                                                                                                                                                                                                                                                                                                                              15-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                          20-0CT-1997;
                                                                                                                                                                                                                                                                                                                                                               29-APR-1999
                                                                                                                                                                                            AAY09373;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                             protein
                                                                                                                                                      g
                                                                                                     qq
           셤
                                    õ
                                                                                 ò
```

```
This present sequence comprises a 42 kDa C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against Plasmodium falciparum. The C-terminal end of the sequence is modified to include a 6xHis tag. Nucleic acids (see AAX25586 and AAX25593) encoding MSP-1-42 have been modified according to a method of the invention in order to improve expression in mammalian cells and in transgenic animals. The invention provides modified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             modified recombinant nucleic acid sequences useful for producing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           recombinant nucleic acid sequences and methods for increasing the mRNA levels and protein expression of proteins that are difficult to express in cell culture systems. especially mammalian cell culture systems or in transgenic animals. The preferred difficult protein candidates for expression are those derived from lower organisms such as parasites, bacteria and viruses that have DNA coding sequences of high Ar content or which have mRNA instability motifs or rare codons relative to the recombinant expression system
299 GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCS 354
                                                 MSP-1; merozoite surface protein; malaria; vaccine; protein engineering; protein expression; codon usage;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Merozoite surface protein MSP-1-42.
                                                                                                                                                                                                                                                                               AAY05833 standard; Protein; 361 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Disclosure; Fig 2; 43pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98US-0085649.
97US-0062592.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98WO-US22225.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          malarial DNA vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 1999-302742/25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       transgenic animal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Meade H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-PSDB; AAX25587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           W09920766-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20-OCT-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                            02-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     29-APR-1999
                                                                                                                                                                                                                                                                                                                                                              AAY05833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Chen LH,
                                                                                                                                                                                                         9
                                                                                                                                                                                                                                      AAY05833
                                                                                                                                                                                                                                                                                                                                                              A PART OF THE PROPERTY OF THE
```

'n

Gaps

DB 20; Length 361;

Pred. No. 1.5e-118;

84.7%; Score 1665.5; 91.0%; Pred. No. 1.5e 7; Mismatches

Conservative

Best Local Similarity Matches 324; Conser

g

à

ò

Query Match

61 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120

1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE 60  Ŋ

```
Modified malarial protein for use in anti-malarial vaccines
                                                              Example; Fig 11; 35pp; English
                         1999-288313/24
          Meade H;
                                 P-PSDB; AAX56008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY05834;
          Chen LH,
                                                                                                                                                    Sequence
                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Key
Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Peptide
                                                                                                                                                                                                                                                                                                                                                                                                        œ
                         WPI;
                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                              AAY05834
   ð
                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                       ä
                                                                                                                                                                                 238
                                                                                                                                                                                                 240
                                                                                                                                                                                                                298
                                                                                                                   NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                           GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                  180
                                                                                                                                                                                                                               299
                                                                                      9
                                                                                             Gaps
                                                                                     1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE
                                                                                                                                                         LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                         VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                        to be used. The invention allows expression of MSP-1 in the milk of transgenic animals, and also provides a DNA vaccine comprising vector containing the altered MSP-1-42 nucleic acid.
                                                                                                                                                                                                                                            /note= "Asn in native MSP-1-42 (N262Q mutation)"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    'note= "Asn in native MSP-1-42 (N181Q mutation)'
                                                       DB 20; Length 361;
                                                                      ж
;
                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                MSP-1; merozoite surface protein; malaria; vaccine; protein engineering; protein expression; codon usage; transgenic animal; mutant.
                                                                                                                                                                                                                                                                                                                                                                                                                                     peptide
                                                                       22;
                                                      Score 1665.5; DB 2
Pred. No. 1.5e-118;
7; Mismatches 22;
                                                                                                                                                                                                                                                                                                                                                 Modified merozoite surface protein MSP-1-42.
                                                                                                                                                                                                                                                                                                                                                                                                                                     signal
                                                                                                                                                                                                                                                                                                                                                                                                                                    /note= "beta-casein
                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                                                                                                                                   AAY09374 standard; Protein; 376 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note= "6xHis tag"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GENZ ) GENZYME TRANSGENICS CORP.
                                                     Query Match 84.7%;
Best Local Similarity 91.0%;
Matches 324; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98US-0085649.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US22226
                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                            371..376
                                                                                                                                                                                                                                                                                                                                                                                               Plasmodium falciparum
                                361 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO9920774-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15-MAY-1998;
20-OCT-1997;
                                                                                                                                                                                                                                                                                                                                  31-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20-OCT-1998;
                                                                                                                                                                                                                                                                                                                                                                                                      Synthetic.
                                                                                                                                                                                                                                                                                                                 AAY09374;
                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                       Key
Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                            Peptide
                                                                                                                   61
                                                                                                                                  61
                                                                                                                                                  121
                                                                                                                                                                 121
                                                                                                                                                                                181
                                                                                                                                                                                                181
                                                                                                                                                                                                                239
                                                                                                                                                                                                                                                                                    양업
ប្បប្បវន្តន
                                                                                                   g
                                                                                                                   ò
                                                                                                                                 g
                                                                                                                                                  ò
                                                                                                                                                                             ŏ
                                                                                                                                                             g
                                                                                                                                                                                              용
                                                                                                                                                                                                               ð
                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                             ò
```

```
GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                ಭ
                                                                                                                                                                                                                                                                                                                                                                                                                                       The present sequence represents a 42 kba C-terminal portion of manalaria marozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against plasmodium falciparum. The sequence has been modified to include an N-terminal beta-casein signal peptide and an N-terminal faxHis and intending on N181Q and N262Q mutations have been introduced eliminate 2 N-91ycosylation sittes. These modifications allow the MSP-1-42 protein to be expressed in the milk of transgenic mice.
                                                                                                                                                                                                                                                                                                                                                                                                               1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               239 VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              299 GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCS 354
                                                                                                                                                                                                                                                                                                                                                             ä
                                                                                                                                                                                                                                                                                                           DB 20; Length 376;
                                                                                                                                                                                                                                                                                                                                                             24; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MSP-1; merozoite surface protein; malaria; vaccine; protein engineering; protein expression; codon usage; transgenic animal; mutant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1..15
/note= "beta-casein signal peptide"
                                                                                                                                                                                                                                                                                                                                   Pred. No. 4.5e-118;
6; Mismatches 24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MSP-1-42.
                                                                                                                                                                                                                                                                                                           Score 1659.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Modified merozoite surface protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note= "6xHis tag"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY05834 standard; Protein; 376
                                                                                                                                                                                                                                                                                                         84.4%;
90.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                           Matches 323; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                   Best Local Similarity
                                                                                                                                                                                                                               376 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                02-AUG-1999
```

ä

us-09-500-376-4.rag

ဖ

Page

```
;
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The present sequence represents a 42 kDa C-terminal portion of malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an important target for the development of a vaccine against plasmodium falciparum. The sequence has been modified to include no N-terminal beta-casein signal peptide and an N-terminal 6xHis tag. In addition, N1810 and N2620 mutations have been introduced to eliminate 2 N-glycosylation sites. These modifications allow the MSP-1-42 protein to be expressed in the milk of transgenic mice. The invention generally relates to modified recombinant nucleic acid sequences and methods for increasing the mRNA levels and protein expression of proteins that are difficult to express in call culture systems. The preferred difficult protein candidates for transgenic animals. The preferred difficult protein candidates for transgenic animals. The preferred difficult protein candidates for expression are those derived from lower organisms such as parasites, bacteria and viruses that have DNA coding sequences of high AT content or which have mRNA instability motifs or rare codons relative to the recombinant expression system to be used.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VFENLLKSILSNLLDWKLARYVKHFTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQE 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181 LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New modified recombinant nucleic acid sequences useful for producing malarial DNA vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVNVKDILNSPFNKRE 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note= "Asn in native MSP-1-42 (N2620 mutation)"
                              'note= "Asn in native MSP-1-42 (N181Q mutation)"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 20; Length 376;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 1659.5; DB 20; Lengtl
Pred. No. 4.5e-118;
6; Mismatches 24; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GENZ ) GENZYME TRANSGENICS CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Example; Fig 11; 43pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    / Match 84.4%;
Local Similarity 90.7%;
hes 323; Conservative (
                                                                                                                                                                                                                                                                                                   98WO-US22225
                                                                                                                                                                                                                                                                                                                                                                    98US-0085649
                                                                                                                                                                                                                                                                                                                                                                                               97US-0062592
   Misc-difference 197
                                                             Misc-difference 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1999-302742/25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Chen LH, Meade H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         376 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAX25593
                                                                                                                                                               WO9920766-A2
                                                                                                                                                                                                                                                                                                20-0CT-1998;
                                                                                                                                                                                                                                                                                                                                                                                            20-0CT-1997;
                                                                                                                                                                                                                                                                                                                                                                15-MAY-1998;
                                                                                                                                                                                                                               29-APR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        239
$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\frac{1}{2}$$\fra
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
```

```
The present invention relates to non-natural variants of a C-terminal fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The non-natural variants have reduced affinity for at least 1 antibody capable of blocking a second antibody that inhibits the proteolytic cleavage of Plasmodium MSP-1_4_2, and has the same affinity for at least one third antibody that inhibits the proteolytic cleavage of plasmodium MSP-1_4_2, and has the non-natural variants of the present invention are useful for immunising a mammal against malaria, and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               can be used to treat malaria. The present sequence is MSP-133 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 NFKNVLESDLIPYKDLTSSNYVVKDPYKFLNKEKRDKFLSSYNYIKDSIDTDINFANDVL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121 GYYKILSEKYKSDLDSIKKYINDKQGENEKYLPFLNNIETLYKTVNDKIDLFVIHLEAKV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LNYTYEKSNVE--IKELIYLKTIQDKLADFKKNNNFVGIADLSTDYNHNNLLTKFLSTGM 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Uthaipibull C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 AVTPSVIHNILSKIENEYEVLYLKPLAGVYRSLKKOLENNVMTFNVNVKDILNSPFNKRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Novel variants of the C-terminal fragment of Plasmodium merozoite surface protein-1, useful as vaccines for treating or preventing malaria
315 gdkcvenpnptcnennggcdadakcteedsgsngkkitcectkpdsyplfdgifcs 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 262;
                                                                                                                                                                                                     Merozoite surface protein; protazoacide; vaccine; malaria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Syed S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Match 63.3%; Score 1246; DB 22;
Local Similarity 95.7%; Pred. No. 6.3e-87;
les 247; Conservative 2; Mismatches 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Morgan W,
                                                                         AAB37610 standard; Protein; 262 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Holder A, Birdsall B, Feeney J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Example 5; Fig 15; 126pp; English.
                                                                                                                                                                      Merozoite surface protein-133
                                                                                                                                                                                                                                                                                                                                                      99GB-0009072.
99US-0311817.
99CA-2271451.
                                                                                                                                                                                                                                                                                                                                                                                                                     (MEDI-) MEDICAL RES COUNCIL.
                                                                                                                                                                                                                                                                                                                           20-APR-2000; 2000WO-GB01558
                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 247; Conservative
                                                                                                                                                                                                                                 Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2001-015762/02.
N-PSDB; AAC68978.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            262 AA;
                                                                                                                                                                                                                                                              WO200063245-A2.
                                                                                                                                                                                                                                                                                                                                                          20-APR-1999;
13-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                      25-MAY-1999;
                                                                                                                                        27-FEB-2001
                                                                                                                                                                                                                                                                                              26-OCT-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                        AAB37610;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                            AAB37610
q
                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
```

÷

GSKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCS 354

299

ò

```
96..116
/note= "glycosylphosphatidylinositol anchoring sequence"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This is the amino acid sequence of a recombinant protein comprising amino acids 1613-1705 of the Plasmodium falciparum merozoite surface amino acids 1613-1705 of the Plasmodium falciparum merozoite surface protein 1 (MSP1) 19 kD C-terminal fragment (p19), linked to a glycosylphosphatidylinositol membrane anchoring sequence. p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the p19 fragment provides a high level of protective immunity since it includes epitopes not presented in the p42 fragment.
                                                                                                                                                                                                                                                                                                                                                                                                        /note= "amino acids derived from P. falciparum MSPl p19
fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 270 KTMIQQNSGCFRHLDEREECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSG 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Recombinant protein containing Plasmodium merozoite surface protein-1 p42 fragment - useful in antimalarial vaccines, also new antibodies for diagnosis and protein purification
                                                                                                                                                                                                                                                    Plasmodium vivax; merozoite surface protein; MSP1; p19;
Plasmodium falciparum; malaria; vaccine; immunity; epitope.
      71 sngkkitcectkpdsyplfdgifcsssnflgisfllilmlilysfi 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               330 SNGKKITCQCTKPDSYPLSMVIFCSSSNFLGISFLLILMLILYSFI 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mendis K, Nato F;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 528; DB 18;
Pred. No. 6.1e-33;
2; Mismatches 8;
                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                  AAW22592 standard; Protein; 116 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Disclosure; Fig 1B; 85pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Barnwell JW, Longacre-Andre S,
                                                                                                                                                                                                                PfMSP1(p19)A protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26.8%;
90.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  97WO-FR00291.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (INSP ) INST PASTEUR.
(UYNY ) UNIV NEW YORK STATE.
                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 90.6
Matches 96; Conservative
                                                                                                                                                                                                                                                                                                            Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1997-425034/39.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              116 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            P-PSDB; AAW22592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO9730159-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                14-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        14-FEB-1996;
                                                                                                                                                                          25-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-AUG-1997
                                                                                                                                                                                                                                                                                                                                  Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                      AAW22592;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Roth C;
                                                                                                                                                                                                                                                                                                                                                                      Key
Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Region
                                                                              AAW22592
                                                                                                                                      용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note= "glycosylphosphatidylinositol anchoring sequence"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This is the amino acid sequence of a recombinant protein comprising amino acids 1613-1705 of the Plasmodium falciparum merozoite surface protein 1 (MSP1) 19 KD C-terminal fragment (p19), linked to a glycosylphosphatidylinositol membrane anchoring sequence. p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the p19 fragment provides a high level of protective immunity since it includes epitopes not presented in the p42 fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                          /note= "amino acids derived from P. falciparum MSP1 p19 fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   270 KTMIQQNSGCFRHLDEREECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSG 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 kkqcpensgcfrhldereeckcllnykqegdkcvenpnptcnennggcdadakcteedsg 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Recombinant protein containing the merozoite surface protein-1 p19 fragment - useful in anti-malarial vaccines, diagnosis and protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 116;
                                                                                                                                                                                                                                                                                      Plasmodium vivax; merozoite surface protein; MSP1; p19;
Plasmodium falciparum; malaria; vaccine; immunity; epitope.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 528; DB 18; Length 1
Pred. No. 6.1e-33;
2; Mismatches 8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nato F;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mendis K,
                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                      ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Disclosure; Fig 1B; 85pp; French.
                                                                                                                                  AAW36103 standard; Protein; 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Longacre-Andre S,
                                                                                                                                                                                                                                                    PfMSP1(p19)A protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 26.8%;
Best Local Similarity 90.6%;
Matches 96; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97WO-FR00290.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NEW YORK STATE.
                    96FR-0001822
239 VFENLLKSILSNLLDWKL 256
                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                96..116
                                                                                                                                                                                                                                                                                                                                              Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PASTEUR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1997-425033/39.
N-PSDB; AAT94550.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (INSP ) INST
(UYNY ) UNIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Barnwell JW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO9730158-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    14-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              purification
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          14-FEB-1996;
                                                                                                                                                                                                              25-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21-AUG-1997
```

Roth C;

Synthetic

Region

Region

AAW36103;

AAW36103

g

ò

ö

Gaps

ö

Indels

g

SNGKKITCQCTKPDSYPLSMVIFCSSSNFLGISFLLILMLLLYSFI 375

Sequence

ð 셤 ò

Length 116;

ပ်

Uthaipibull

Syed S,

AAB37608;

12

AAB37608 RESULT

```
The present invention relates to non-natural variants of a C-terminal fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The non-natural variants have reduced affinity for at least 1 antibody capable of blocking a second antibody that inhibits the proteolytic cleavage of Plasmodium MSP-1_4_2, and has the same affinity for at least one third antibody that inhibits the proteolytic cleavage of plasmodium MSP-1_4_2, and has the non-natural variants of the present invention are useful for immunising a mammal against malaria, and can be used to treat malaria. The present sequence is MSP-119 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            270 KTMIQQNSGCFRHLDEREECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSG 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Novel variants of the C-terminal fragment of Plasmodium merozoite surface protein-1, useful as vaccines for treating or preventing malaria .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Chimeric; Plasmodium vivax; merozoite surface protein; MSP1; p19; Plasmodium falciparum; malaria; vaccine; immunity; epitope.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22.7%; Score 447; DB 22; Length 108; 89.8%; Pred. No. 7.7e-27; ive 1; Mismatches 8; Indels
                                                       Merozoite surface protein; protazoacide; vaccine; malaria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Morgan W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               330 SNGKKITCQCTKPDSYPLSMVIFCSSSN 357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers 1.19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAW22593 standard; Protein; 127 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Holder A, Birdsall B, Feeney J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Example 5; Fig 15; 126pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          - Plasmodium vivax.
- Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PfMSP1(p19)S protein sequence.
Merozoite surface protein-119
                                                                                                                                                                                                                                                                                                                                                       99GB-0009072.
99US-0311817.
99CA-2271451.
                                                                                                                                                                                                                                                                                                20-APR-2000; 2000WO-GB01558.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (MEDI-) MEDICAL RES COUNCIL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25-MAR-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity 89.8
nes 79; Conservative
                                                                                                                 Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2001-015762/02.
N-PSDB; AAC68977.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     108 AA;
                                                                                                                                                                           WO200063245-A2.
                                                                                                                                                                                                                                                                                                                                                                                     13-MAY-1999;
25-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                          20-APR-1999;
                                                                                                                                                                                                                                     26-OCT-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Chimeric Chimeric
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAW22593;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAW22593
NAME OF COLOR OF STREET OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The non-natural variants have reduced affinity for at least 1 antibody capable of blocking a second antibody that inhibits the proteolytic cleavage of Plasmodium MSP-1_4_2, and has the same affinity for at least one third antibody that inhibits the proteolytic cleavage of Plasmodium MSP-1_4_2, and has the same affinity for at least one third antibody that inhibits the proteolytic cleavage of Plasmodium MSP-1_4_2, compared to natural MSP-1_1_9. The present sequence is the wild-type MSP-1 protein. This sequence was used to generate the variants of the present invention. The non-natural variants of the present invention. The non-natural variants of the present invention are useful for immunising a mammal against malaria, and can be
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ပဲ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      270 KTMIQQNSGCFRHLDEREECKCLLNYKQEGSKCVENPNPTCNENNGGCDADAKCTEEDSG 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Uthaipibull
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        merozoite
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Novel variants of the C-terminal fragment of Plasmodium merozoite surface protein-1, useful as vaccines for treating or preventing malaria -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 96;
                                                                                                                                                                                                                                                                                                                           Merozoite surface protein; protazoacide; vaccine; malaria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Syed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 447; DB 22;
Pred. No. 6.6e-27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Morgan W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       330 SNGKKITCQCTKPDSYPLSMVIFCSSSN 357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Example 2; Page 48; 126pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ą.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ٦,
                                                                            AAB37608 standard; protein; 96 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB37609 standard; Protein; 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Feeney
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22.7%;
89.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99GB-0009072.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0311817.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (MEDI-) MEDICAL RES COUNCIL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20-APR-2000; 2000WO-GB01558
                                                                                                                                                                                                                                                                    Merozoite surface protein-1
                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 89.8 ses 79; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Holder A, Birdsall B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          used to treat malaria.
                                                                                                                                                                                                                                                                                                                                                                                  Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2001-015762/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               96 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200063245-A2
                                                                                                                                                                                                    27-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-MAY-1999;
25-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                27-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 26-OCT-2000
```

Sequence

Query Match

Matches

AAB37609;

AAB37609 ID AAB3 XX AC AAB3 XX DT 27-F

RESULT 13

g

g ŏ

ò

ö

Gaps

```
Chimeric
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                               1;
                                                                                                                                                                                                                                                                                                                           This is the amino acid sequence of a chimeric protein comprising amino acids 1-32 of the plasmodium vivax merozoite surface protein 1 (MSP1) linked to amino acids Asn1613-Ser1705 which correspond to the 19 kD C-terminal fragment (p19) of MSP-1 from Plasmodium falciparum. p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the p19 fragment provides a high level of protective immunity since it includes epitopes not presented in the p42 fragment.
                                                        t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              7; Gaps
                                             33..34
/note= "encoded by restriction enzyme sequence used
create the chimeric sequence"
                                                                                                                                                                                                                                                                       Recombinant protein containing Plasmodium merozoite surface protein-1 p42 fragment - useful in antimalarial vaccines, also new antibodies for diagnosis and protein purification
                                                                                /note= "derived from P. falciparum C-terminal p19
fragment of MSP1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Chimeric; Plasmodium vivax; merozoite surface protein; MSP1; p19; Plasmodium falciparum; malaria; vaccine; immunity; epitope.
                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 127;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   305 NPNPTCNENNGGCDADAKCTEEDSGSNGKKITCQCTKPDSYPLSMVIFCS 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              22.0%; Score 432.5; DB 18; Length 70.0%; Pred. No. 1.2e-25; ive 12; Mismatches 14; Indels
                                    /note- "derived from P. vivax MSP1"
                                                                                                                                                                                                                 Mendis K, Nato
       20..127 ... /note= "mature protein"
/note= "signal peptide"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ā
                                                                                                                                                                                                                                                                                                            Disclosure; Fig 1C; 85pp; French.
                                                                                                                                                                                                                  s,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAW36102 standard; Protein; 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PfMSP1(p19)S protein sequence.
                                                                                                                                                                                                                  Longacre-Andre
                                                                                                                                                 97WO-FR00291
                                                                                                                                                                                              NEW YORK STATE.
                                                                                                                                                                   96FR-0001821
                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 70.09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                     PASTEUR
                                                                                                                                                                                                                                            WPI; 1997-425034/39.
P-PSDB; AAW22592.
                                                                                                                                                                                                                                                                                                                                                                                                                127 AA;
                                                                                                                                                                                     (INSP ) INST
(UYNY ) UNIV
                                                                                                                                                                                                                 Barnwell JW,
                                                                                                            WO9730159-A2
                                                                                                                                                 14-FEB-1997;
                                                                                                                                                                   14-FEB-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25-MAR-1998
                                                                                                                               21-AUG-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAW36102;
                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
          Protein
                                                                                                                                                                                                                           Roth C;
                            Region
                                              Region
                                                                         Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAW36102
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

```
Ξ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This is the amino acid sequence of a chimeric protein comprising amino acids 1-32 of the Plasmodium vivax merozoite surface protein 1 (MSP1) linked to amino acids Asn1613-Ser1705 which correspond to the 19 KD C-terminal fragment (p19) of MSP-1 from Plasmodium falciparum. p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the p19 fragment provides a high level of protective
                                                                                                                                                                                                                                                 ಭ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     immunity since it includes epitopes not presented in the p42 fragment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      245 KSILSNLLDWKLARYVKHPTTPMRKKTMIQQNSGCFRHLDEREECKCLLNYKQEGSKCVE 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25 kqlvanvdefnisqh-----qcvkkqcpensgcfrhldereeckcllnykqegdkcve 77
                                                                                                                                                                                                                                             Recombinant protein containing the merozoite surface protein-1 p19 fragment - useful in anti-malarial vaccines, diagnosis and protein purification
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7;
                                                                                                                                                                                                                                                                                                           /note= "derived from P. falciparum C-terminal fragment of MSP1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               22.0%; Score 432.5; DB 18; Length 127; 70.0%; Pred. No. 1.2e-25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         14; Indels
                                                                                                                                                                                                   /note= "derived from P. vivax MSP1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mendis K, Nato
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           12; Mismatches
                                                                                                         /note= "signal peptide"
                                                                                                                                                     /note= "mature protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8, 2001, 12:31:56
                                                                     Location/Qualifiers
Plasmodium vivax.
Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Disclosure; Fig 1C; 85pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Barnwell JW, Longacre-Andre S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     97WO-FR00290.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 96FR-0001822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INST PASTEUR.
UNIV NEW YORK STATE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Search completed: August
Job time: 164 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1997-425033/39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
ses 77; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ¥;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAT94549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 127
                                                                                                                                                                                                                                                                                                                                                                           WO9730158-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     14-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 14-FEB-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INSP ) INST
                                                                                                                                                                                                                                                                                                                                                                                                                         21-AUG-1997.
      1 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local &
                         Chimeric
                                                                     Key
Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (UXNX)
                                                                                                                                   Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Roth C;
                                                                                                                                                                               Region
                                                                                                                                                                                                                          Region
                                                                                                                                                                                                                                                                                           Region
```

## THIS PAGE BLANK (USPTO)

plasmodium plasmodium plasmodium

plasmodium plasmodium plasmodium

plasmodium plasmodium plasmodium plasmodium

plasmodium

plasmodium plasmodium

025724 025721 025725 025725 025723 025723 025723 025727 025717 025717 025717 025717 02572 02572 02572 02572 02572 02572 02572 02572 02572 02572 02572 02572 02572 02572

Sequence:

ö

Searched:

Database

plasmodium

plasmodium plasmodium plasmodium plasmodium

Lasmodium

```
1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRRY 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN=RO-71;
MEDLINE-9275047; PubMed=1592091;
MEDLINE-9275047; PubMed=1592091;
01afsson P., Matile H., Certa U.;
"Plasmodium falciparum: the repetitive MSA-1 surface protein of the RO-71 isolate is recognized by mouse antibody against the nonrepetitive repeat block of RO-33.";
Exp. Parasitol. 74:381-389(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Tolle R., Bujard H., Cooper J.A.; "Plasmodium falciparum: variations within the C-terminal region of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEROZOITE SURFACE ANTIGEN 1 (FRAGMENT).
Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 5; Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1
74134 MW; AA2137E699255150 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99.1%; Score 2043.5; DB 5;
ilarity 99.5%; Pred. No. 2.2e-102;
Conservative 1; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     651 AA
                                                                                                                                                                                                                                                                                                                                                                                                                           ALIGNMENTS
 025724
025721
043995
025725
043996
                                                                           Q25723
Q25726
O43997
                                                                                                                     025722
025727
025728
025728
                                                                                                                                                                                             025719
025720
025720
025981
003999
025966
025976
025984
025976
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                025973
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=95354793; PubMed=7628566;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              merozoite surface antigen-1.";
Fxp Parasitol. 81:47-54(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-MAY-2000 (TrEMBLrel. 13,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Exp. Parasitol. 81:47-54(1
EMBL; 235329; CAA84558.1;
InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 392; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         651 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
   NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-RO-71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Merozoite.
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   025924
025924;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
Q25924
   ACCOOR REPORTED THE REPORT OF 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                      (without alignments)
743.944 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 plasmodium
plasmodium
plasmodium
plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             plasmodium
plasmodium
plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  plasmodium
plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q25924 plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       asmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          asmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     asmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                           1 AISVTMDNILSGFENEYDVI.....SNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
                                                                                                       ; Search time 70.07 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q25983
Q25969
Q25974
Q25977
Q25977
Q25977
Q91215
Q91215
Q91214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       09ty92
025923
025923
025978
025970
025980
025980
               GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                               Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                      425026 segs, 132305027 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                         protein – protein search, using sw model
                                                                                                     August 8, 2001, 12:29:12
                                                                                                                                                                                                                                          Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q25969
Q25974
Q25975
Q25977
Q25979
Q9T245
Q9T247
Q9T274
Q25968
Q9T274
Q25968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  025967
025970
025980
025982
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_unclassified:*
sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                               Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_rodent:*
                                                                                                                                                                  us-09-500-376-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sp_plant:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                        SPTREMBL_16:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_fungi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    569
569
569
569
1694
1704
570
570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_mhc:*
                                                                                                                                                                                                                           BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2043.5
2036.5
2030.5
2030.5
2030.5
2030.5
2030.5
2027
2027
2027
2027
2021.5
2011.5
2011.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score
                                                                                                                                                                               Perfect score:
                                                                                                                                                                                                                           Scoring table:
```

ä

1;

Result

```
241 LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                     301 HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    YNNLYNKIDDYLINLKAKINDCNYEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            241 LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        237 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    416 LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             416 LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
356 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                        l part of the precursor to
of Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                 Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5; Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A26C45914AF01396 CRC64;
                                                                                                                                                                                                                                                                            01-NOV-1996 (TrEMBLrel. 01, Created)
NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURPACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 2030.5; DB 5;
Pred. No. 9.7e-102;
0; Mismatches 2;
                                                                                                                                                                                                                                                    Ş
                                                                                                                                                           536 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                            361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                 569
                                                                                                                                                                                                                                                                                                                                                                                                                                       Jongwutiwes S., Tanabe K., Kanbara H.; "Sequence conservation in the C-terminal major merozoite surface proteins (MSPI) of field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mol. Biochem. Parasitol. 59:95-100(1993)
EMBL; D13346; BAA02607.1; -.
                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE=93295445; PubMed=8515786;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64485 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98.5%;
ilarity 99.2%;
Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR000561; -. Pfam; PF00008; EGF; 1. Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       569 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
Matches 391; Conserv
                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NON_TER
SEQUENCE
                                                                                                                                                                                                                                                 025969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               361
                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                   025969
                        οy
                                                   g
                                                                                    ò
                                                                                                              g
                                                                                                                                            ò
                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     major merozoite surface proteins (MSP1) of Plasmodium falciparum from fileld isolates."; urface proteins (MSP1) of Plasmodium falciparum from fileld isolates."; bardies (MSP1) of Plasmodium falciparum from Mol. Blochem. Parasitol. 59:95-100(1993).

EMBL; D13355; BAA02616.1; -.

Pfam; PF00008; EGF; 1.

Merozoite; EGF-11ke domain.
                                                                                     YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                 437
                                                                                                                                            YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKALDDKIDLFKNHNDFEAIKK 240
                                                                                                                                                                                                      300
                                                                                                                                                                                                                                  557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240
                                                                                                                                                                                                                     YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK
                           FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS
                                                                                                   LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK
                                                                                                                                                                                                                                                                              AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                          FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 AISVTADNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                57A6B62FF72CE885 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
MSP1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. 4.6e-102;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                              ¥
                                                                                                                                                                                                                                                                                                                                        Query Match 98.8%; Score 2036.5; Best Local Similarity 99.5%; Pred. No. 4.6e Matches 392; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                         PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                            569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
MEDLINE-93295445; PubMed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
"Sequence conservation in the C-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               569 AA; 64491 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               177
                         61
                                                       319
                                                                                    121
                                                                                                               379
                                                                                                                                            181
                                                                                                                                                                         438
                                                                                                                                                                                                                                498
                                                                                                                                                                                                                                                               301
                                                                                                                                                                                                                                                                                            558
                                                                                                                                                                                                                                                                                                                       361
                                                                                                                                                                                                                                                                                                                                                    618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                   8
                                                                                                         g
                                                                                                                                         ò
                                                                                                                                                                   g
                                                                                                                                                                                                   à
                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                        à
                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
```

ï

ಭ

535

025974; 025974

용

```
Jongwutiwes S., Tanabe K., Kanbara H.; "Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEOKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           301 HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                 241 LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.

MEDLINE=91295445; PubMed-8515786;
Jongwutiwes S., Tanabe K., Kanbara H.;
Sequence S., Tanabe K., Kanbara H.;
"Sequence of the precursor to major merozoite surface proteins (MSP1) of Plasmodium falciparum field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                               1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK
                                                                                                                                                                                                                                                                                                                            Length 569;
                   Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                          ED4284B2867C9703 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
MASP1.
                                                                                                                                                                                                                                                                                                                          Score 2030.5; DB 5;
Pred. No. 9.7e-102;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 536 PDSYPLEDGIFCSSSNFLGISFLLILMLILYSFI 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    269
                                                                                                                                                                   Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13349; BAA02610.1; -
InterPro; IPR000561; -
Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mol. Biochem. Parasitol. 59:95-100(1993)
EMBL; D13350; BAA02611.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                     PubMed=8515786;
                                                                                                                                                                                                                                                                        569 AA; 64459 MW;
                                                                                                                                                                                                                                                                                                                        Query Match 98.5%;
Best Local Similarity 99.2%;
Matches 391; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
   Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Plasmodium falciparum.
                                                                   SEQUENCE FROM N.A.
                                                                                     MEDLINE=93295445;
                                 NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota;
                                                                                                                                                                                                                                                          NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  025977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  025977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ဖ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
025977
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                             Ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .
1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               416 LINDDTKKDWLGKLLSTGLVQNFPWTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-93295445; PubMed-8515786; Jongwutiwes S., Tanabe K., Kanbara H.; Sequence conservation in the C-terminal part of the precursor to major merozoite surface proteins (MSP1) of Plasmodium falciparum field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    569;
                                                                                                                                                                                                     Plasmodium falcipārum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A2691FE569E80396 CRC64;
                                                                                                                            01-MAY-2000 (TrEMBLrel. 01, Last sequence update) MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT) MSP1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01, Created)
01, Last sequence update)
13, Last annotation update)
PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 2030.5; DB 5 Pred. No. 9.7e-102;
                                                                                     ΑĄ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    536 PDSYPLFDGIFCSSSNFLGISFLLIIMLILYSFI 569
536 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         569
                                                                                   569
                                                                                                                                                                                                                                                                                                                                                                        Mol. Blochem. Parasitol. 59:95-100(1993)
EMBL; D13348; BAA02609.1;
                                                                                                                   Created)
                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               64457 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98.5%;
99.2%;
                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR000561; -... Pfam; PF00008; EGF; 1. Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel. 0:
01-NOV-1996 (TrEMBLrel. 0:
01-MAY-2000 (TrEMBLrel. 1:
MAJOR MEROZOITE SURFACE PI
MSP1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 391; Conservative
                                                                                                                   (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               569 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                      NCBI_TaxID=5833;
                                                                                                              01-NOV-1996 (
01-NOV-1996 (
01-MAY-2000 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                               NON_TER
SEQUENCE
```

Query Match

8

õ g ò g ò 025975 025975;

DT 011

241

ò

g

g

301

ò 셤 à 셤

Gaps 9

1;

236

415

300 475 535

RAM FE SO

ద ò g ò a à

```
1362 FLDVLESDLAGFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 1421
                                                                                                                                                                                                                           HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                  416 LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241 LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL
                                                                                     FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS
                                                                                                                                                                                               YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL
                                                                                                                                                                                                                                                                                                        YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Jiang G., Liu R.Z., Daubenberger C.A., Pluschke G.; Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases. EMBL; AF062348; AAC72884.1; -. InterPro; IRR000561; -. Pfam; PF00008; EGF; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1694 AA; 192794 MW; 84CFC0E709F5673B CRC64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAY-2000 (TrEMBLrel. 13, Created)
1-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 2030.5; DB Pred. No. 3e-101;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 1694 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  536 PDSYPLEDGIFCSSSNFLGISFLLILMLILYSFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98.5%;
99.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 98.5
Best Local Similarity 99.2
Matches 391; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-HN1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Merozoite.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9TZT5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               241
                                                                                     61
                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9TZT5
                                                                                                                                      g
                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ID
DDT TO
DDT TO
OOS OOS
OOS NEED
DDS N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                  Ω
                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                        ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1;
                                                                                                                                                                                                                                                                                                                                                                                                                                               415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-93295445; PubMed-8515786; Jongwutiwes S., Tanabe K., Kanbara H.; "Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates.";
                                                                                                                                                                                                                                                                                                                                  61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        297 YYEKVLAKYKDDLESIKKVIKEEKE-FPSSPPTTPPSPAKTDEOKKESKFLPFLTNIETL 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 AISVTMDNILLSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                  1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1;
                                                                                                                                                                                                  569;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length
                                                                                                                                                                                                  Length
                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A2691FE569E80396 CRC64;
                                                                                                            A26C45914AF01396 CRC64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ou-nov-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
MSP1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 2030.5; DB 5;
Pred. No. 9.7e-102;
0; Mismatches 2; I
                                                                                                                                                                                               DB 5;
                                                                                                                                                                                            Score 2030.5; DB 5;
Pred. No. 9.7e-102;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               536 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mol. Blochem. Parasitol. 59:95-100(1993).
EMBL; D13352; BAA02613.1; -.
Pfam; PF00008; EGF; -.
Merozoite; EGF-like.domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  64457 MW;
                                                                                                            64485 MW;
                                                                                                                                                                                               98.5%;
99.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Match 98.5%;
Local Similarity 99.2%;
des 391; Conservative
                         Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                            Query Match 98.5
Best Local Similarity 99.2
Matches 391; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Plasmodium falciparum
InterPro; IPR000561;
                                                                                                            569 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  569 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-5833;
                                                                               NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
```

025979; 025979

RESULT
7025979
10 20379
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001
10 001

241

ò

g

셤

301

ŏ

g

g

ò

1;

Gaps 09

ä

Indels

Length 1694;

240

Matches

ò

Ą.

1704

S

1541

ð g Q9NHX1

RESULT Q9NHX1

1661

g

à

```
SEQUENCE FROM N.A. MEDLINE=93295445; PubMed=8515786;
                                                                                                                                                                                                                                                                                                      1704 AA; 193762 MW;
                                                                        01-MAY-2000 (TrEMBLrel. 13,
01-JUN-2000 (TrEMBLrel. 14,
MEROZOITE SURFACE PROTEIN 1
                                                                                                                                                                                                                                                                                                                                                 98.5%;
99.2%;
                                                       (TrEMBLrel. 13, (TrEMBLrel. 13, (TrEMBLrel. 14,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                             391; Conservative
                              PRELIMINARY;
                                                                                                                                                                                                                                                                        Pfam; PF00008; EGF; 1.
                                                                                                                              Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                            NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=5833;
                                           Q9TZT4;
01-MAY-2000 (
01-MAY-2000 (
01-JUN-2000 (
                                                                                                                                                                                                     STRAIN-HN2;
                                                                                                                                              Eukaryota;
                                                                                                                                                                                                                                                                                         Merozoite
                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           025968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            025968
                              Q9TZT4
 10
                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  301
 RESULT
                D9TZT4
                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AC DATE OF THE STATE OF T
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1540
1601 HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 1660
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                        Shan Z.X., Yu X.B., Li X.R., Ma C.L., Fang J.M.;
"Molecular cloning and sequence analysis of major merozoite surface antigen(gpl95)gene of plasmodium falciparum isolate FCCI/HN.";
submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AF218248; AAF27526.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK
                                          301 HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 1694;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1;
                                                                                                                                                                                                                                                                                                              Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 2030.5; DB 5; Length
Pred. No. 3e-101;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        B51634A49E0F6728 CRC64;
                                                                                                                                                                                                                                            Last sequence update)
Last annotation update)
                                                                                                                  PDSYPLEDGIFCSSSNFLGISFLLILMLILYSFI 1694
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                 361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                     PRT; 1694
                                                                                                                                                                                                                                Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1694 AA; 192766 MW;
                                                                                                                                                                                                                                            01-OCT-2000 (TIEMBLIEL 15, Last 01-MAR-2001 (TIEMBLIEL 16, Last MAJOR MEROZOITE SURFACE ANTIGEN GP195.
                                                                                                                                                                                                                           (TrEMBLrel. 15, C
(TrEMBLrel. 15, 1
(TrEMBLrel. 16, 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               / Match 98.5%;
Local Similarity 99.2%;
hes 391; Conservative
                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR003247; -. Pfam; PF00008; EGF; 1. ProDom; PD001527; -; 1.
                                                                                                                                                                                                                                                                                                     Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                             STRAIN-FCC1/HN;
                                                                                                                                                                                                                 Q9NHX1;
01-OCT-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Merozoite.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
```

Matches

ð g ò g ò 8 ò

1362 121 1422 181 241 1541

ογ

g

Ω

301 1601 361

ò

q

ò

1661

```
61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1704;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                   Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                 Jiang G., Liu R.Z., Daubenberger C.A., Pluschke G.;
Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
EMBL; AF062249; AAG72885.1; -.
InterPro; IPR000561; -.
                                                                                                                                                                                                                                                                                                                                                                                          385526D0DA56FD1D CRC64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                     Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 2030.5; DB 5;
Pred. No. 3.1e-101;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1671 PDSXPLFDGIFCSSNFLGISFLLILMLILXSFI 1704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          570 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Jongwutiwes S., Tanabe K., Kanbara H.;
                                                              PRECURSOR
Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
```

ij

ဖ

```
Q25922
Q25922;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Tolle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         225922
               SO FT
                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDIJNE-88011243; PubMed-3079521;
Tanabe K., Mackay M., Goman M., Scaife J.G.;
"Allelic dimorphism in a surface antigen gene of the malaria parasite
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-93295445; PubMed-8515786; Jongwutiwes S., Tanabe K., Kanbara H.; Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   536
          art of the precursor to the
Plasmodium falciparum from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                   1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     241 LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                              177 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                Length 570;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                      424BF553CCC2F2BE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-UIN-2000 (TrEMBLrel. 14, Last annotation update)
MAJOR MEROZOITE SUFFACE PROTEIN PRECURSOR (FRACMENT).
                                                                                                                                                                                                                                                                                                                Score 2027; DB 5;
Pred. No. 1.5e-101;
"Sequence conservation in the C-terminal part major merozoite surface proteins (MSP1) of Pla field isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         537 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            570 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                  1; Mismatches
                                                                            Mol. Biochem. Parasitol. 59:95-100(1993).
EMBL; D13345; BAA02606.1; -.
InterPro: PR000561; -.
Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mol. Biochem. Parasitol. 59:95-100(1993)
EMBL; D13343; BAA02604.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Plasmodium falciparum.";
J. Mol. Biol. 195:273-287(1987).
                                                                                                                                                                                                                                    570 AA; 64632 MW;
                                                                                                                                                                                                                                                                                                                98.3%;
98.7%;
                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 98.7
Matches 389; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=5833;
                                                                                                                                                                                                            NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9TYG2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
1097462
       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pan W., Tolle R., Bujard H.;
"A direct and rapid sequencing strategy for the Plasmodium falciparum antigen gene gp190/MSA1.";
antigen gene gp190/MSA1.";
Bubl. Biochem. Parasitol. 7:244(1995).
EMBL: 235327; CAA84556.1;
InterPro: IRRO00561;
Pfam: PF00008; EGF; 1.
                                                                                                                                                                                                                                                                                177 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY 236
                                                                                                                                                                                                                                                                                                                                   121 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                       181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 360
                                                                                                                                                                                                                                                              61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                         1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY 60
                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
PRECURSON FOR THE MAJOR MEROZOLTE SURFACE ANTIGENS.
Plasmodium falciparum (isolate NF54).
Eukaryota. Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TAXID=5843;
                                                                                                                    Length 570;
                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Submitted (JUL-1994) to the EMBL/GenBank/DDBJ databases.
                                                                   8674DEC89B2D662A CRC64;
                                                                                                                  Query Match 98.3%; Score 2027; DB 5; Best Local Similarity 98.7%; Pred. No. 1.5e-101; Matches 389; Conservative 1; Mismatches 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 1720 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Tolle R., Bujard H., Cooper J.A.;
Exp. Parasitol. 0:0-0(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=96123395; PubMed=8577332;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1069-1720 FROM N.A.
                                                                 64630 MW;
              Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
IPR000561;
                                                                 570 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN=NF54;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN=NF54;
                                                  NON_TER
SEQUENCE
```

ö

```
1
652 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 389; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                569 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
        Merozoite.
                      NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              025978
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  025978
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15
                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
Q25978
        FT SO
                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                  ö
                                                                                                                                                                                                                                                    "Variation in the precursor to the major merozoite surface antigens of Plasmodium falciparum.";
Mol. Biochem. Parasitol. 27:291-302(1988).
                                                                                                                                                                        1327 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY 1386
                                                                                                                                                                                                                                                                                                                                                                                                       300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              360
                                                                                                                                                                                                                                FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                    Gaps
                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKK1TCECTK
                                                                                                                                                      1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                     181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                               LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-1907/93; PubMed-7628566;
TOLLE N., Bujard H., Cooper J.A.;
"Plasmodium falciparum: variations within the C-terminal region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Peterson G.M., Coppel R.L., McIntyre P., Langford C.J., Woodrow Brown G.V., Anders R.F., Kemp D.J.;
                                                                                                                  ö
                                                                              Length 1720;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                  Indels
  POTENTIAL.
W; 717B2FD1E637A8A3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Submitted (JUL-1994) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Last sequence update)
Last annotation update)
                                                                      Score 2025; DB 5;
Pred. No. 6.1e-101;
1; Mismatches 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 1720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-MAY-2000 (TrEMBLrel. 13, Last annota MEROZOITE SURFACE ANTIGEN 1 (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN=FCH5/NF7;
MEDLINE-88142999; PubMed=2449612;
                  195725 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          merozoite surface antigen-1.";
Exp. Parasitol. 81:47-54(1995).
EMBL; 235328; CAA84557.1; -.
InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
                                                                            98.2%;
98.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1996 (TrEMBLrel. 01, 01-MAY-2000 (TrEMBLrel. 13.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1996 (TrEMBLrel. 01,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 1-298 FROM N.A.
                                                                    Query Match
Best Local Similarity 98.7
Matches 389; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                      ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
  1720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-FCH5/NF7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-FCH5/NF7;
SIGNAL
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Tolle R.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      025923
                                                                                                                                                                                                                                                                     1387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                               241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    361
                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25923

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010

2010
  SO
                                                                                                                                                      ò
                                                                                                                                                                                        ద
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                   Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
```

```
Ή
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=93295445; PubMed=8515786; Jongwutiwes S., Tanabe K., Kanbara H.; Jongwutiwes S., Tanabe K., Kanbara H.; "Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSPI) of Plasmodium falciparum from field isolates.";
                                                                                                                                                                                                                                                                       360
                                                                                                                                                                                                                                                 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                    YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                 1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                                                                                       241 LINDDTKKDMLGKLLSTGLVQNFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY
                                                                                                                     ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           569;
                                                                                  Length 652;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBL_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length
                                                                                                                       Indels
1
74292 MW; 2B6A87737B490A62 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            64536 MW; B8B9B63EECB1DA51 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
MAJOR MEROZOITE SURFACE PROTEIN PRECURSOR (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5,
                                                                              Score 2021; DB 5;
Pred. No. 3.6e-101;
2; Mismatches 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 2017.5; DB 5,
Pred. No. 4.8e-101;
1; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           619 PDSYPLEDGIFCSSSNFLGISFLLILMLILYSFI 652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mol. Biochem. Parasitol. 59:95-100(1993)
EMBL, D13351; BAA02612.1; -.
Interpro; IPR000561; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                              98.0%;
98.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         97.8%;
98.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1996 (TrEMBLrel. 01,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF00008; EGF; 1.
Merozoite; EGF-like domain.
                                                                              Query Match 98.0°
Best Local Similarity 98.5°
Matches 388; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
```

```
181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNHNDFEAIKK 240
                                                          61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                               361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                      PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 569
                                                                                                                                             536
           g
                                    qq
                                                               g
                                                                                          q
                                                                                                       ò
                                                                                                                   a
                                                                                                                                ö
                                                                                                                                          qq
õ
                         δ
                                                   ò
                                                                             ò
```

Search completed: August 8, 2001, 12:33:42 Job time: 270 sec

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

Run on:

August 8, 2001, 12:32:29; Search time 24.79 Seconds (without alignments) 544.440 Million cell updates/sec

Title: Perfect score: Sequence:

US-09-500-376-3. 2052 1 AISVYMDNILSGFENEYDVI......SNFLGISFLLILMLILYSFI 394

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

93435 segs, 34255486 residues Searched:

93435 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_39:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	819		PO4934 plasmodium	P50495 plasmodium							-			P32380 saccharomyc	P27895 saccharomyc		Q00798 plasmodium	P25386 saccharomyc	P12753 saccharomyc	P16154 clostridium	P36112 saccharomyc.	Q11208 sarcophaga	P05556 homo sapien	P43573 saccharomyc			_	~		_		466	P47166 saccharomyc
SUMMARIES	ΩI	S	MSP1_PLAFM	MSP1_PLAFC	MSP1_PLAFP	MSP1_PLAF3	MSP1_PLAFK	MSP1_PLAFW	MSP1_PLAYO	YL17_CAEEL	YIO9_YEAST	MYS2_DICDI	ECM_HUMAN	MLP1_YEAST	NUF1_YEAST	CIN8_YEAST	SMC2_YEAST	RBP1_PLAVB	USO1_YEAST	RA50_YEAST	TOXA_CLODI	YKZ6_YEAST	PPOL_SARPE	ITB1_HUMAN	YFC3_YEAST	MYS2_SCHPO	HMW2_MYCGE	ABRA_PLAFC	KAR3_YEAST	YAE7_YEAST	NOF_DROME	ABRA_PLAFG	RBP2_PLAVB	YJ9C_YEAST
	DB	-	٦	Н	-	Н	٦	П	٦	٦	Н	-	-	٦	٦	7	-	7	7	٦	Н	٦	٦	٦	٦	П	П	1	٦	7	-	Н	П	1
	Length	1701	1701	1726	1726	1682	1630	1639	1772	1130	1679	2116	1228	1875	944	1038	1170	2869	1790	1312	2710	540	966	198	196	1526	1805	743	729	622	984	900	1251	707
de	Query		99.1	97.8	97.6	96.5	61.1	61.1	28.9	7.2	7.2	7.1	6.5	6.5	6.4	•	•	6.3	٠	•	6.2	•		6.2			٠	•	•	•	•	5.9	•	5.9
	Score	2033	2033	2006	2003	1980.5	1254.5	1254.5	593.5	147	147	145.5	133	132.5	132	130.5	130	130	129	128.5	127.5	127	127	126.5	126	125.5	125.5	124.5	124	123.5	123.5	122	122	121.5
	Result No.	Ц.,,		Э	4	5	9	7	80	σ	10	11	12	13	14	15	. 16	17	18	19	20	21	22	23	24	25	56	27	28	58	30	31	32	33

P08964 saccharomyc . Q00402 saccharomyc	013766 schizosacch Q10411 schizosacch P06101 saccharomyc	P47907 penicillium Q9zl21 helicobacte	P47037 saccharomyc P54697 dictyosteli	P41467 autographa P75093 mycoplasma	051578 borrelia bu
MYS1_YEAST NUM1_YEAST	YE94_SCHPO YD86_SCHPO CC37_YEAST	RMS5_PENUR UVRC_HELPJ	SMC3_YEAST MYSJ_DICDI	Y066_NPVAC Y018_MYCPN	EX5B_BORBU
					-
1928 2748	512 1957 506	399 594	1230 2245	808 1030	1169
5.9	ი. ი. ი. ი. ი. ი.	55.00 58.00	5.8 8.8	7. 7. 8. 8.	5.8
121.5	121 121 120.5	120	120 120	119 119	119
34 35	36 37 38	39 40	4142	4 4 4	45

## ALIGNMENTS

NDARD; PRT; 1701 AA. 13, Created) 13, Last sequence update) 34, Last annotation update) PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)	modium falciparum (isolate FC27 / Papua New Guinea).  ryota; Alveolata; Apicomplexa: Haemosporida; Plasmodium.  TaxID=8837;  ENCE FROM N.A.  IND=8814299; PubMed=2449612;  IND=8814299; PubMed=2449612;  IND=8814299; PubMed=20.1;  IND=8014299; PubMed=20.1;  Ind=10.1;  IND=10.2;	yright. It is produced through e of Bloinformatics and the El Institute. There are no rest. utions as long as its content is not removed. Usage by an agreement (See http://www.isb::	tein; Repeat; Signal; POTENTIAL. MEROZOITE SURFACE P N-LINKED (GLCNAC
	MASPOLITE AND CAPE AN	entry is copentry is copentry is copentric to informatics of it institutes statement is a license to license	EMBL; M19143; AAA29653.1; - PIR; A54498; A54498. InterPro; IPR005651; - Ffam; PF00008; EGF; 1. Malaria; Merozaite; Polypro Fransmembrane; GPI-anchor. SIGNAL 1 19 CARBOHYD 239 239 CARBOHYD 110 110 CARBOHYD 470 CARBOHYD 536 536 CARBOHYD 536 536 CARBOHYD 607 607 CARBOHYD 899 899 CARBOHYD 899 899 CARBOHYD 899 899 CARBOHYD 919 919
0 (Re 0 (Re 6 (Re SURFA	nodium falci ryota; Alvec TaxID=5837; ENCE FROM N. IRE-8814299; IRE-8814299; FROM M. G.V. A.A.C n G.V. A.A.C lacinon in the lacinon in the lacinon in the lacinon in the lacinon in the lacinon in the lacinon in the lacinon in the lacinon in the lacinon in the lacinon in the lacinon in the lacinon in the lacinon in the la	SWISS-PROT en the Swiss Buropean Bioin by non-profiled and this les requires and an email the swiss but the swiss	143; AN 1PR000 1PR000 1PR000 10008; C 20 110 239 470 470 607 802 802 802 803 919 919 965
_PLAFF MSPl_PLAFF P13819; 01-JAN-1990 01-JAN-1990 01-CCT-1996 MEROZOITE SI (PMMSA).	MSF-1.  MSF-1.  MSF-1.  Bukaryota; Alveola  Bukaryota; Alveola  (1)  SEQUENCE FROM N.A.  MEDLINE-88142999;  Peterson M.G., Cop  Brown G.V., Anders  "Variation in the  for Plasmodium falco  for Plasmodium falco  for Plasmodium falco  in Subcellular LO  (POTENTIAL).	This SWISS-PROT ent between the Swiss the European Bioinf use by non-profit modified and this s entities requires a or send an email to	EMBL: M19143; AAA29653.1; PIR; A54498; A54498. InterPro; IPR000561; Pfam; PF00008; EGF; 1. Malaria; Merozoite; Polyp Transmembrane; GFI-anchor SIGNAL 10 110 CARBOHYD 239 239 CARBOHYD 239 239 CARBOHYD 239 239 CARBOHYD 536 536 CARBOHYD 607 607 CARBOHYD 607 607 CARBOHYD 607 607 CARBOHYD 607 607 CARBOHYD 802 899 CARBOHYD 899 899 CARBOHYD 919 919 CARBOHYD 965 965 CARBOHYD 965 965
MSP1_ 1D AC DT DT DT DE		388888888	PI P

```
361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MSP1_PLAFC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1668
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                     ö
                                                                                    240
                                                                                                                                                                                                                                                  LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                             HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-88011243; PubMed-3079521;
Tanabe K., Mackay M., Goman M., Scaife J.G.;
"Allelic dimorphism in a surface antigen gene of the malaria parasite
                                                                                                                    FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                               YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 42
                                                      Gaps
                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                         YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK
                                                                         1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                              30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OMEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U., Stunnenberg H., Bujard H.; Polymorphism of the precursor for the major surface antigens of Plasmodium falciparum merozoites: studies at the genetic level."; EMBO J. ♣:3823-3829(1985).
                                                     ö
                               Length 1701;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Plasmodium falciparum (isolate mad20 / Papua New Guinea)
3920B75E73D38552 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Tanabe K.;
Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
                              Score 2033; DB 1;
Pred. No. 2.7e-97;
0; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                  PDSYPLEDGIFCSSNFLGISFLLIALLIXSFI 1701
                                                                                                                                                                                                                                                                                                                                                                                                                  Ş
                                                                                                                                                                                                                                                                                                                                        PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                  1701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-86136024; PubMed-3004972;
193719 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REVISIONS TO 1403; 1569 AND 1629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Plasmodium falciparum.";
J. Mol. Biol. 195:273-287(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                    01-AUG-1988 (Rel. 08, Created)
                              Ouery Match 99.1%;
Best Local Similarity 99.2%;
Matches 391; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 1-115 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-70153;
1701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (PMMSA) (P190).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                  MSP1_PLAFM
P08569;
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Tanabe K
                                                                                                                                                                                                                                                                                                                                                                                           RESULT 2
MSP1_PLAFM
                                                                                                                                                              121
                                                                                                                                                                                   1428
                                                                                                                                                                                                                             1488
                                                                                                                                                                                                                                                                        1548
                                                                                                                                                                                                                                                                                                                   1608
                                                                                                                    61
                                                                                                                                        1368
                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                  241
                                                                                                                                                                                                                                                                                             301
                                                                                                                                                                                                                                                                                                                                       361
g
                                                                                           g
                                                                                                                   9
0
0
                                                                                                                                                                               g
                                                                                                                                                                                                       ò
                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                 Op
                                                                                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                            ò
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation—the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1308 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNINLNDILNSRLKKRKY 1367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1368 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNTLLKSYKXIKESVENDIKFAQEGIS 1427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           301 HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0
                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                  InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 2033; DB 1; Length 1701; Pred. No. 2.7e-97;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC. . .) (POT
MW; 3FC2EC59AF96EA98 CRC64;
                                                                                                                                                                                                                                                                                            MEROZOITE SURFACE PI
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (Rel. 10, Created)
(Rel. 10, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. 2.76
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI
                                                                                                                                                                                                                                                                                POTENTIAL
                                                                                                                                              EMBL; X05624; CAA29112.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 99.1%;
Best Local Similarity 99.2%;
Matches 391; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 193768
                                                                                                                                                                                                                                                                 GPI-ancho
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                100
1100
1100
1110
2339
603
803
803
913
965
11196
11388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AA;
                                                                                                                                                                 PIR; A26868; A26868.
PIR; B25120; B25120.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         991
1089
1196
1588
                                                                                                                                                                                                                                                               Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13-AUG-1987
01-MAR-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MSP1_PLAFC
P04934;
                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                  SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TASE
Tase
```

m

```
1453 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 1512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "Plasmodium falciparum: gene structure and hydropathy profile of the major merozoite surface antigen (gpl95) of the Uganda-Palo Alto
                                                                                                                                                                                                                                                                                                            01-OCT-1996 (Rel. 34, Created)
1-OCT-1996 (Rel. 34, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                            181 YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK
                                          LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                     HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-89005525; PubMed-3049134;
Chang S.P., Kramer K.J., Yamaga K.M., Kato A., Case S.E.,
                                                                                                                                                                                                                                                                                                                                                                                             Plasmodium falciparum (isolate Palo Alto / Uganda)
                                                                                                                                                                                                       1693 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 1726
                                                                                                                                                                                       361 PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=57270;
                                                                                                                                                                                                                                                                                                                                                                  (PMMSA) (GP195).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Siddiqui W.A.;
                                                                                                                                                                                                                                                                                  MSP1_PLAFP
P50495;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              isolate.",
                                                                                                                                                                                                                                                         RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    major
                                                      1513
                                                                                                          1573
                                                                                                                                                             1633
                                                                                 241
                                                                                                                                     301
                                                                                                                                                                                                                                                                                    g
                                                                                                                                                             g
                                                                                                                                                                                                            qq
                          Ω
                                                                               οχ
                                                                                                     g
                                                                                                                                   ò
                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                         Weber J.L., Sim B.K.L., Lyon J.A., Wolff R.;
"Merozoite surface protein sequence from the Camp strain of the human malaria parasite Plasmodium falciparum.",
Nucleic Acids Res. 16:1206-1206(1988).
-i- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                 οĘ
                                                                                                                                                                                                                                                                                                         PTM: MEROZÓITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 4
KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                      SEQUENCE OF 1-1103 FROM N.A.
MEDLINE-86205236; PubMed-3517809;
Meber J.L., Lehninger W.M., Lyon J.A.;
Wariation in the gene encoding a major merozoite surface antigen the human malaria parasite Plasmodium falciparum.";
Nucleic Acids Res. 14:3311-3323(1986).
01-OCT-1996 (Rel. 34, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            pfam. PF00008; EGF; 1.
Malaría; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROTEIN 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DD8AD45FA352BCF3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEROZOITE SURFACE PI
N-LINKED (GLCNAC.
                                                   Plasmodium falciparum (isolate Camp / Malaysia)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -LINKED (
                                                                                                                                                                                                   SEQUENCE OF 1104-1726 FROM N.A. MEDLINE-88143999; PubMed=3278296;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; X03831; CAA27446.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            196197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Transmembrane; GPI-anchor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR000561;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SAZOGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1613
1726
                                                                                                                                                                                                                                                                                           (POTENTIAL)
                          (PMMSA) (P195).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIGNAL
```

PRT; 1726 AA

STANDARD;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                               KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
Parasitol. 67:1-11(1988).
SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                          (POTENTIAL).
-!- PIM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                       . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                     Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                  SURFACE PROTEIN 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROZOITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                      EMBL; M37213; AAA29611.1; -.
                                                                                                                                                                                                                                                                                                                                                                                          GPI-anchor
                                                                                                                                                                                                                                                                                                                                                                                                                                                   133
                                                                                                                                                                                                                                                                                                                           InterPro; IPR000561;
Pfam; PF00008; EGF; 1
                                                                                                                                                                                                                                                                                                                                                                                          Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                             SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
```

ö

Gaps

; 0

7; Indels

Length 1726;

Score 2006; DB 1; Pred. No. 6.6e-96;

97.8%; 98.0%;

Best\_Local Similarity 98.0 Matches 386; Conservative

Query Match

1; Mismatches

1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY

YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180

FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120

61 1393 121

· 8 ·δ 8 ò

```
EMBL; M35727; AAA29715.1; -. EMBL; Y00087; CAA68280.1; -. EMBL; Z35326; CAA84555.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Match 96.5%;
Local Similarity 97.2%;
es 383; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            192462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                   IPR000561;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AA;
                                                                                                                                                                                       PIR; S06286; S06286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13-AUG-1987 (Rel. 0
01-FEB-1996 (Rel. 3
01-OCT-1996 (Rel. 3
MEROZOITE SURFACE P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (PMMSA) (P190).
                                                                                                                                                                                                                                                                       Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MSP1_PLAFK
P04932;
                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                  TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                        SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 6
MSP1_PLAFK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          301
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BEALAGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                               ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                            240
                                                                                                                                                                                                                                                                                                                                                     YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                     FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-88166657; PubMed-3327688;
Certa U., Rotmann D., Matile H., Reber-Liske R.;
An naturally occurring gene encoding the major surface antigen
Precursor p190 of Plasandium falciparum lacks tripeptide repeats.";
EMBO J. 6:4137-4142(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PTM: MEROZOTTE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 4 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  merozoite surface antigen-1.";
Exp. Parasitol. 81:47-54(1995).
-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                     1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                   YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SURFACE ANTIGENS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1032-1682 FROM N.A.
MEDLINE-9358-7959, PubMed-7628566;
Tolle R., Bujard H., Cooper J.A.;
"Plasmodium falciparum: variations within the C-terminal region
N-LINKED (GLCNAC. . .) (POTENTIAL)
N-LINKED (GLCNAC. . .) (POTENTIAL)
MW; 5B59CEEFA2F9A026 CRC64;
                                                                                                                                               ;
0
                                                                                                       Length 1726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Plasmodium falciparum (isolate ro-33 / Ghana).
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBL_TaxID=5834;
                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MSP1_PLAF3 STANDARD; PRT; 1682 AA. P19588: Q25921; Ol-FEB-1991 (Rel. 17, Created) 01-NOV-1997 (Rel. 35, Last sequence update) 01-NOV-1997 (Rel. 35, Last annotation update) MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE :
                                                                                                     Score 2003; DB 1; Pred. No. 9.4e-96; 2; Mismatches 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                     97.68;
97.78;
1221 1221
1613 1613
1726 AA; 196174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 1-1061 FROM N.A.
                                                                                                                        Local Similarity 97.7 ses 385; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (PMMSA) (P190).
CARBOHYD
                       CARBOHYD
                                        SEQUENCE
                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MSPL PLAF3
ID MSPL PLAF3
DT 01-FEBB
DT 01-FEBB
DE MENOVO
DE MENOVO
DE MENOVO
DE MENOVO
ON NCBL TY
RP SEQUEN
RP SEQUEN
RP A DAT
RP FECUT
RP SEQUEN
RP SEQUEN
RP FECUT
RP SEQUEN
RP SEQUEN
RP FECUT
RP SECUT
RP SECU
                                                                                                                                                                                                                                                                                                           1393
                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          361
                                                                                                                                               Matches
                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1693
                                                                                                                                                                                                                                                                                                                                                                                                                                 181
  FT
                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                   ç
                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
```

```
matics Institute. There are no restrictions on its institutions as long as its content is in no way tement is not removed. Usage by and for commercial
                                                                                        entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1290 AISVTMDNILSGFENEYDVIXLKPLAGVYRSLKKQIEKNIFTFNLNLNDILNSRLKKRKY 1349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1589 HLDEREECKCLLNYKQEGDKCVENPPTCNENNGGCDADAKCTEEDSGSNGKKITCECTK 1648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121 YYEKVLAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 FLDVLESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGIS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               05, Created)
33, Last sequence update)
34, Last annotation update)
PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             YNNLVNKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKK
                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
MEMBRANE ANCHOR (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 AISVTMDNILSGFENEYDVIYLKPLAGVYRSLKKOIEKNIITFNLNLINDILNSRLKKRKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Length 1682;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ij
                                                                                                                                                                                                                                                                                                                                     Pfan; PP00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INKED (GLCNAC. . . C82A1E159948CAD6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 1980.5; DB 1;
Pred. No. 1.3e-94;
2; Mismatches 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1649 PDSYPLFDGIFCSSSNFLGISFLILLMLILYSFI 1682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 1630 AA
                               use by non-profit institutions as long modified and this statement is not removed.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PDSYPLFDGIFCSSSNFLGISFLLILMLILYSFI
```

S

```
A24594; A24594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=5848;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                (PMMSA) (P195).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Freeman R.R.;
                                                                                                                                                                                                                                                                                                                                                 MSP1_PLAFW
P04933:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REVISIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIGNAL
                     1421
                                                                                                                                   303
                                                                                                                                                                                                               363
                                                                                                                                                                                                                                                                                                                               MSP1_PLAFW
                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                   HERE HERE HERE WAS DRAWNED OF COCCOCCAR BARKER BARK
                                                           δ
                                                                                             g
                                                                                                                                   οy
                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit. institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .;
m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                186 NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                         PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 4 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                               Pan W., Tolle R., Bujard H.;
Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6 MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVL
                                                                                       SEQUENCE FROM N.A.
MEDLINE=86136024; PubMed=3004972;
Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,
Stunnenberg H., Bujard H.;
"Polymorphism of the precursor for the major surface antigens of
Plasmodium falciparum merozoites: studies at the genetic level."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61.1%; Score 1254.5; DB 1; Length 1630; 61.2%; Pred. No. 2.1e-57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           . 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR000561; -.
Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                 falciparum (isolate Kl / Thailand).
Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           73; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROTEIN 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADBDEC3CE0A46322 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEROZOITE SURFACE
TRIPEPTIDE SG(TP)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED (GLCNAC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEMBRANE ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         54; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (
                                                                                                                                                                                                                                                 REVISIONS, SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; X03371; CAA27070.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             187289
                                                                                                                                                                                                           EMBO J. 4:3823-3829(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GPI-anchor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIR; A25120; SAZQK1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                     NCBI_TaxID=5839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1156
1165
1436
                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Fransmembrane;
               Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Sim
Matches 240;
                                       Eukaryota;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                MEDLINE-86014355; Pubmed=2995820;
Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V.,
Nicholls S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-AUG-1987 (Rel. 05, Created)
10-FEB-1996 (Rel. 33, Last sequence update)
01-FEB-1996 (Rel. 34, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Submitted (MAR-1991) to the EMBL/GenBank/DDBJ databases.
                                                                                  246 TKKDMLGKLLSTGLVQIFPN---TIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHL
                                                                                                                                                                                                DEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKK1TCECTKPD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            structure of the precursor to the three major surface of plasmodium falciparum merozoites.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Plasmodium falciparum (isolate Wellcome).
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL.
MENOZOTTE SURPACE PEN-LINKED (CLCUAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 1639 AA
                                                                                                                                                                                                                                                                                                                                        1599 SYPLFDGIFCSSSNFLGISFLLILMLILYSFI 1630
                                                                                                                                                                                                                                                                                                             SYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; X02919; CAA26676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "Primary structure of the antigens of Plasmodium fal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Nature 317:270-273(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GPI-anchor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR000561; -.
```

ဖ

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1489
                                                                                                                                                                                                                                                                                                                                302
                                                                                                                                                                                                                                                                                                                                                                                                               66 ESDLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKV 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        126 LAKYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDD 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCECTKPD 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Burns J.M. Jr., Daly T.M., Vaidya A.B., Long C.A.;
"The 3' portion of the gene for a Plasmodium yoelii merozoite surface antigen encodes the epitope recognized by a protective monoclonal antibody.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence analysis upstream of the gene encoding the precursor to the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        42
                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                         MDNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVL 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PTM: MEROZÓITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 4
KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1430 DKIDLFVIHLEAKVLNYTYEKSNVEVKIKELNYLKTIQDKLADFKKNNNFVGIADLSTDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         246 TKKDMLGKLLSTGLVQIFPN---TIISKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-JAN-1990 (Rel. 13, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                               (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                      25;
                                                                                                                                                                                                   Length 1639;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                     73; Indels
  N-LINKED (GLCNAC. ..) (POTE N-LINKED (GLCNAC. ..) (POTE N-LINKED (GLCNAC. ..) (POTE N-LINKED (GLCNAC. ..) (POTE MW; 2C255B6616C87F6E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            or merozoite surface antigens of Plasmodium yoelii.";
. Biochem. Parasitol. 39:285-288(1990).
                                                                                                                                                                                                DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Proc. Natl. Acad. Sci. U.S.A. 85:602-606(1988).
                                                                                                                                                                                             $; Score 1254.5; DB

; Pred. No. 2.1e-57
54; Mismatches 7:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 1772 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SYPLEDGIFCSSSNFLGISFLLILMLILYSFI 1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           363 SYPLFDGIFCSSSNFLGISFLLILMLILYSFI 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE=90205979; PubMed=2320061;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=88124889; PubMed=2448778;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 1093-1772 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JAN-1990 (Rel. 13, Created)
01-JAN-1990 (Rel. 13, Last sequ
1165
1174
1445
1526
AA; 187618 M
                                                                                                                                                                                             61.1%;
61.2%;
                                                                                                                                                                                     Query Match
Best Local Similarity 61.2<sup>3</sup>
Matches 240; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Plasmodium berghei yoelii.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (PMMSA) (230 KDA).
  1165
1174
1445
1526
1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Alveo
NCBI_TaxID=5862;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-17XL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MSP1_PLAYO
P13828;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Lewis A.P.;
                                                                               CARBOHYD
SEQUENCE
  CARBOHYD
                            CARBOHYD
                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MSPI_PLAYO

ID MSPI_PLAYO

MSPI_PLAYO

MSPI_PLAYO

DT 01-0AN

DT 01-0AN

DT 01-0AN

DE (PMMSA)

DE (PM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  303
                                                                                                                                                                                                                                                                                                         9
  FT
                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        :| || :|:||:| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1521 LYNTQLAAVKEQIATIE------AETNDTNKEEKKKYIPILEDLKGLYETVI 1566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1567 GOAEEYSEELONRLDNYKNEKAEFEILTKNLEKYIQIDEKLDEFVEHAENNKHIASI--A 1624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KYKDDLESIKKVIKEEKEKFPSSPPTTPPSPAKTDEQKKE--SKFLPFLTNIETLYNNLV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              68 DLMQFKHISSNEYIIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKVLA 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             33; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8 NILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVLES 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NKIDDYLINLKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLF----KNTNDFEAIKKL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              242 INDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNIS-QHQCV-KKQCPENSGCF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              300 RHLDEREECKCLLNYKQ-EGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKKITCEC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 1772;
                                                                                                                                                                                                                                                                                                                                                      Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -> V (IN REF. 2).
9A6291658EB0F45D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                            PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28.9%; Score 593.5; DB 1; 32.5%; Pred. No. 1.8e-23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          96; Mismatches 137;
                                                                                                                                                                                                                                                                                                                                                                                                                 MEROZOTTE SURFACE PI
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             359 TKPDSYPLFDGIFCSSSNFLGISFLLILMLILYS 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 1130
                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                 EMBL; J03612; AAA29762.1; -. EMBL; J04668; AAA29702.1; -. PIR; A28121; A28121. PIR; A45532.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA; 197230
                                                                                                                                                                                                                                                                                                                                                                             GPI-anchor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 128; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               646
829
1018
1090
1408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1541
1629
1680
1521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       35,
35,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (Rel.
(Rel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1446
1541
1629
1680
1521
                                                                                                                                                                                                                                                                                                                                                                          Transmembrane;
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   YL17_CAEEL
Q11102;
01-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-NOV-1997
15-JUL-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 9
YL17_CAEEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              186
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DI PI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Д
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
```

```
S0001411; MLP2.
                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                            Query Match
Best Local Simi
Matches 105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MYS2_DICDI
P08799;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MYS2_DICD]
 Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ٠Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                        14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               81 IIEDSFKLLNSEQKN-ILLKSYKYIKESVENDIKFAQEGISYYEKVLAKYKDDLESIKKV 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IKEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLVNKIDDYLINLKAKI 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               863
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         714 IAQLNEKI-----STKETEKDSELEKTIAQLEIDNSSKSDQIEKLHLRVNDML 761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     307
                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                    Rhabditida; Rhabditoidea;
                                                                                                                                                                                                                                                                                                                                                                                                                                             SLKKQIEKNIITFNLNDILNSRLKK - - - - - - - RKYFLDVLESDLMQFKHISSN - - EY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKL--INDDTKKDMLGKLLST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          762 DOMGTIKDEL---VKKNEEIKTISAKTAQLLESNTVESETKLASVTEEREKE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     258 GLVQIFPNTIISKLIEGKFQDMLNISQ-----HQCVKKQCPENSGCFRHLDE----REE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
01-FEB-1995 (Rel. 31, Last annotation update)
HYPOTHETICAL 195.1 KDA PROTEIN IN DNA43-UBII INTERGENIC REGION
                                                                                                                                                                                                                                                                                                                                                                       Score 147; DB 1; Length 1130; Pred. No. 0.9;
                                                                                                                                                                                                                                                                    EMBL; U41545; AAA83190.1; -.
WormPep; C02F12.7; CE03901.

Hypothetical protein; Colled coil.

DOMAIN 121 779 COLLED COIL (POTENTIAL).

DOMAIN 805 1061 COLLED COIL (POTENTIAL).

SEQUENCE 1130 AA; 131485 MW; BOFDZEFE3D99FB09 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                     64; Mismatches 129; Indels
                                                                                                              Miller N.; Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
HYPOTHETICAL 131.5 KDA PROTEIN C02F12.7 IN CHROMOSOME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ¥
                                    Eukaryota; Metazoa; Nematoda; Chromadorea;
Rhabditidae; Peloderinae; Caenorhabditis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Saccharomyces cerevisiae (Baker's yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1679
                                                                                                                                                                                                                                                                                                                                                                            7.2%; Score 147; 23.6%; Pred. No. 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        308 CKCLLNYKQEGDKCEENPNPTCNENN 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TTVVLFDNSIQEKIDEK-EATINEMN 888
                                                                                                                                          -! - SIMILARITY: WEAK, TO MYOSINS
                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity 23.6 nes 77; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                     Caenorhabditis elegans
                                                                                     SEQUENCE FROM N.A. STRAIN-BRISTOL N2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-FEB-1995 (Rel
01-FEB-1995 (Rel
01-FEB-1995 (Rel
                                                            NCBI_TaxID=6239;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              YIO9_YEAST
P40457;
                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
YIO9_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 864
                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                               31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RNX OCC SEE DIT SEE
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DTYRTLMEEIESLKKELQIFKTANSSSDAFEKLKVNMEKEKDRIIDERTKEFEKKLQETL 1415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1416 NKSTSSEAEYSKDIETLKKEWLKEYEDETLRRIKEAEENLKKRIRLPSEERIQKIISKRK 1475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 60 YFLDVLESDLMQFKHISSNEY----IIEDSFKLLNSE--QKNILLKSYKYIKESVEN--- 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----DIKFAQEGISYYEKVLAKYKDDLE-SIKKVIKEEKEKFPSSPPTTPPSPAKTD 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EQKKESKFLPFLTNIETLYNNLVNKIDDYLINLKAKI-------NDCNVEKDEAHV 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         212 KI-----TKLSDLKAIDDKI--DL-FKNTNDFEAIKKL------INDDTKK 248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -- CKCLLNYK 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SVTMDNILSGFENEYDVIYLKPLAGVYRSLKKQ---IEKNIITFNLNLNDILNSRLKKRK 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DMLGKLLS-----TGLVQIF------FULVQIF
Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D., Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D., Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 147; DB 1; Length 1679; Pred. No. 1.4;
                                                                                                                                                                  Walsh S.V., Whitehead S.;
Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hypothetical protein.
SEQUENCE 1679 AA: 195141 MW; 298950CC52202D8F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Dictyostelium discoideum (Slime mold).
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
NCBI_TaxID=44689;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   70; Mismatches 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     316 QEGD----KCEENPNPTCNENNGGCDADATCTEEDSGSSRKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-OCT-1989 (Rel. 12, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
MYOSIN II HEAVY CHAIN, NON MUSCLE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NISQH------QCVKKQCPE--NSGCFRHLDEREE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1988 (Rel. 09, Created)
01-OCT-1989 (Rel. 12, Last sequ
15-JUL-1998 (Rel. 36, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; Z38059; CAA86129.1; -. PIR; S48385; S48385.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
```

```
Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Dictyostellum discoldeum myosin motor domain.";
J. Mol. Biol. 274:394-407(1997).

-!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
ACTIVITY THAT IS ACTIVATED BY ACTIN.

-!- SUBBUNT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Machine C.A., Rayment I.,;
Smith: C.A., Rayment I.,;
"X-ray structure of the magnesium(II).ADP.vanadate complex of the
                                                                                                                                                                                                                                                        "Replacement of threonine residues by serine and alanine in a phosphorylatable heavy chain fragment of Dictyostelium myosin II."; FEBS Lett. 269:239-243(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-97452580; PubMed-9305951; Salike A.M., Bauer C.B., Thoden J.B., Rayment I.; Marper C.B., Thoden J.B., Rayment I.; Marray structures of the MgADP, MgATPgammaS, and MgAMPPNP complexes of the Dictyostellum discoideum myosin motor domain."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-98070605; PubMed-9405148;
Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
"X-ray crystal structure and solution fluorescence characterization of Mg.2'(3')-0-(N-methylanthraniloy1) nucleotides bound to the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             of the
                                                                                                                                                                                                                                                                                                                                                                           Wagle G., Noegel A., Scheel J., Gerisch G.; "Phosphorylation of threonine residues on cloned fragments of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    domain of Dictyostelium
                                      MEDLINE-87092266; PubMed-3540939; Wartick H.M., de Lozanne A., Lelinwand L.A., Spudich J.A.; Wartick H.M., de Lozanne A., Lelinwand L.A., Spudich J.A.; Conserved protein domains in a myosin heavy chain gene from Dictyostellum discoideum.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "X-ray structure of the magnesium(II)-pyrophosphate complex truncated head of Dictyostellum discoideum myosin to 2.7-A
                                                                                                                                                                                                                      Lueck-Vielmeter D., Schleicher M., Grabatin B., Wippler J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "X-ray structures of the myosin motor domain of Dicty discoideum complexed with MgADP.BeFx and MgADP.AlF4-. Biochemistry 34:8960-8972(1995).
                                                                                                                    Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762
                                                                                                                                                             PHOSPHORYLATION SITES, AND MUTAGENESIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Dictyostellum discoldeum myosin motor
Biochemistry 35:5404-5417(1996).
                                                                                                                                                                                                                                                                                                                                                                                                               Dictyostellum myosin heavy chain.";
FEBS Lett. 227:71–75(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Biochemistry 36:11619-11628(1997).
                                                                                                                                                                                               MEDLINE-90353583; PubMed-2387408;
                                                                                                                                                                                                                                                                                                                                                        MEDLINE-88112226; PubMed=2828113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-95345066; PubMed-7619795;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-95345067; PubMed-7619796;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-96206189; PubMed-8611530;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ochemistry 34:8973-8981(1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Smith C.A., Rayment I.;
                                                                                                                                                                                                                                                                                                                                      PHOSPHORYLATION SITES.
[1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                        Serisch G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   resolution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Rayment I.;
                                                                                                                                                                               STRAIN-AX2
```

SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL

CORTEX.

DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN BE FURTHER SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institution. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12;
-i- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEFTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
-i- PTM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES THE ACTIN-ACTINATED ATPASE ACTIVITY.
-i- MISCELLANGOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-IP POSITION (688)
                                                                                                               SIMILARIY: THE GLOBULAR HEAD AA SEQUENCE SHOWS A HIGH DEGREE OF SIMILARIY: THE GLOBULAR HEAD SEQUENCES OF MUSCLE & NONWUSCLE HEAVY CHAINS. BY CONTRAST THE ROD SEQUENCE IS LESS CONSERVED, BUT THE PERIODICITIES OF HYDROPHOBIC & CHARGED RESIDUES, WHICH DICTATE THE ALPHA HELICAL COTLED-COLL STRUCTURE, ARE CONSERVED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1673 IKSLVAEVDEVKEQLEDEILA------KDKLVKAKRALEVELEEVRDQLEEEEDSRS 1723
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81 IIEDSFKLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKVLAKYKDDLESIKKVI 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               141 KEEKEKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNLV-----NKIDDYLIN 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       195 LKAKINDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDDTKKDMLGKL 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             22 LKPLAGVYRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDV-LESDLMQFKHISSNEY 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Colled coll, Actin-binding, ATP-binding, 3D-structure, repeat pattern, Methylation, Alkylation, Phosphorylation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RODLIKE TAIL (S2 AND LMM DOMAINS).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 2116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACTIN-BINDING.
ACTIN-BINDING.
METHYLATION (DI.) (POTENTIAL).
ALKYLATION (SH-1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (BY MHCK).
(BY MHCK).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2FC3770BB1EE56A1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PHOSPHORYLATION (BY MHCK).
PHOSPHORYLATION (BY MHCK).
PHOSPHORYLATION (BY MHCK).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COILED COIL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7.1%; Score 145.5; DB:1;
23.1%; Pred. No. 2.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61; Mismatches 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF00612; IQ; 1.
Pfam; PF00063; myosin_head; 1.
PRINTS; PR00193; MYOSINHEAVY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                18 752 0 130 0 130 1823 1833 2029
                                                                                                                                                                                                                                                                                                                                    EMBL; M14628; AAA33227.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PDB; 1LVK; 28-JAN-98.
DictyDb; DD01008; mhcA.
InterPro; IPR000048; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              nterPro; IPR001609; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Heptad repeat pattern;
                                                                                                                                                                                                                                                                                                                                                                                            17-AUG-96.
03-DEC-97.
                                                                                                                                                                                                                                                                                                                                                                                                                                      17-AUG-96.
17-AUG-96.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     23-DEC-96
                                                                                                                                                                                                                                                                                                                                                                                LMMA; 03-DEC-97
                                                                                                                                                                                                                                                                                                                                                                                                                          03-DEC-97
                                                                                                                                                                                                                                                                                                                                                               $00250
                                                                                                                                                                                                                                                                                                                                                   A26655; A26655
S00250; S00250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 82; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        638
738
130
678
1823
1833
2029
                                                                                                                                                                                                                                                                                                                                                                                               MMD;
                                                                                                                                                                                                                                                                                                                                                                                                                          LMMN;
LMND;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1VOM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       1MNE:
                                                                                                                                                                                                                                                                                                                                                                                                              MMG;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOD_RES
MOD_RES
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
DOMAIN
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MOD_RES.
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Myosin;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                     PIR;
PIR;
                                                                                                                                                                                                                                                                                                                                                                                            PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                          PDB;
PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PDB;
                                                                                                                                                                                                                                                                                                                                                                                  PDB;
      á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

σ

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                 MULTIMERS. THE SMALLEST OF WHICH IS A HOWOTRIMER. PROTECLINKED MULTIMERS.

THE PROMULTIMERS THE SMALLEST OF WHICH IS A HOWOTRIMER. PROTECLISIS OF THE PROMULTIMERIN IN THE N-TEMBINAL REGION, LEADS TO THE MATURE PLISS FORM THAT IS STORED IN PLATELETS.

TISSUE SPECIFICITY: SYMTHESIZED BY ENDOTHELIAL CELLS AND MEGAKARYOCYTES. STORED IN PLATELET ALPHA GRAULLES AND ENDOTHELIAL CELL WRIBEL-PALADE BODIES, FOLLOWING ACTIVATION OF THESE CELLS, IT IS RELEASED AND ATTACHED TO MEGAKARYOCYTES, PLATELETS, ENDOTHELIUM AND SUBRUDOTHELIUM OF BLOOD VESSELS. NOT FOUND IN PLASMA. FOUND IN PLASUES SUCH AS PLACEMTA, LUNG, AND LIVER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PTM: EXTENSIVELY N-GLYCOSYLATED.

DISEASE: DEFICIENCY IN MULTIMERIN DUE TO PROTEOLYTIC DEGRADATION
MITHIN THE PLATELET ALPHA GRANULES IS ASSOCIATED WITH AN AUTOSOMAL.

DOMINANT BLEEDING DISORDER (FACTOR V QUEBEC).

SIMILARITY: CONTAINS 1 GIQ DOMAIN.
SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
1876 IDNLRAQIEDEGKIKMRLEKEKRALEGELEELRETVEEAEDSKSEAEQS---KRLVELEL 1932
                                  LSTGLVQIFPNTII----SKLIEGKFQDMLNISQHQCVKKQCPENSGCFRHLDEREE
                                                                                                       1933 EDARRNLQKEIDAKEIAEDAKSNLQREIVEAKGRLEBESIARTNSDRSRKRLEAE 1987
                                                                                                                                                                                                                                                                          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                    308 CKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEE-----DSGSSRKKITCE 357
                                                                                                                                                                                                  15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
39-MAY-12000 (Rel. 39, Last annotation update)
ENDOTHELIAL CELL MULTIMERIN PRECORSOR.
                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A., AND SEQUENCE OF 368-376.
TISSUE-Endothelial cells;
MEDLINE-95355440; PubMed-7629143;
                                                                                                                                                                            PRT; 1228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRINTS; PR00007; COMPLEMNTC1Q.
PROSITE; PS01113; C1Q; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; U27109; AAC52065.1; -.
                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR000152; -. InterPro; IPR000561; -. InterPro; IPR001073; -. Pfam; PF00008; EGF; 1.
                                                                                                                                                                                                                                                                   sapiens (Human)
                                                                                                                                                                                                 15-JUL-1998
15-JUL-1998
30-MAY-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               601456;
                                                                                                                                                                          ECM_HUMAN
                                                                                                                                                                                                                                                                   Ношо
                                                                                                                                                              ECM_HUMAN
                                   ò
                                                           셤
                                                                                   à
                                                                                                        셤
```

```
17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EKFPSSPPTTPPSPAKTDEQKKESKFL-----PFLTNIETLYNNLVNKIDDYLINLKAK 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       199 INDCNVEKDEAHVKITKLSDLKAIDDKIDLFKNTNDFEAIKKLINDDTKKDMLGKLLSTG 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                88 LLNSEQKNILLKSYKYIKES---VENDIKFAQEGISYYEKVLAKYKDDLESIKKVIKEEK 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         80; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29 YRSLK-KQIEKNIITFNLNLNDILNSRLKKRKYFLDVLESDLMQFKHISSNEYIIEDSFK 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ESVVSI-------AAQQKFVLVQENRPTLTDIVELRNHIVNVRQEMTLTCEKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            525 QVSDQKNAPAAAESVSNNVTEYMS-TLHENIKKQSLMMLQMFEDLHIQESKINNLTVSLEM
                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                   CELL ATTACHMENT SITE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1228;
                                                    ENDOTHELIAL CELL MULTIMERIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mismatches 111; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LVQIFPNTISKLIEGKFQDMLNISQHQCVKKQCPENSGCFR--HLDE---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EAA98BOA17E2C4CD CRC64
                                                                                          COILED COIL (POTENTIAL)
COILED COIL (POTENTIAL)
COILED COIL (POTENTIAL)
                                                                                                                                COILED COIL (POTENTIAL)
COILED COIL (POTENTIAL)
COILED COIL (POTENTIAL)
POLY-SER.
                           domain; Coiled coil.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 1;
                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                              BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
N-LINKED (GLCN
N-LINKED (GLCN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6.5%; Score 133; DB 21.5%; Pred. No. 5.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        584 EKESLRGECEDMLSKCRNDFKFQLKDTEEN 613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----REEC----KCLLNYKQEGDKCEEN 324
                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                      N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
N-LINKED
                                                                 EGF-LIKE.
                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                      POLY-GLN
                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                68;
                           EGF-like
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               138071
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                71; Conservative
                                                                                                                                                                                                    188
1056
1065
1076
                           Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
PS00022;
PS01186;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        981
1020
1075
1228
                                                                                                                                                                                                                             DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
            PROSITE;
                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                DISULFID
                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                          Signal;
SIGNAL
                                                                  DOMAIN
                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
```

```
AELLKVVRNLADKLESKEKKSK----QSLQKIESETVNEAKEAIITLKSEKMDLESRIE 601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-94064779; PubMed-7503995;
                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=92176232; PubMed=1541631;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; Z11582; CAA77668.1; -. EMBL; X73297; CAA51733.1; -.
                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
STRAIN-S288C / AB972;
                                    DCNVEKDE----
                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUCLEOLUS
                                                                                                                                                                                                                          NUF1_YEAST
P32380;
                                                                                                                                                                                                   T 14
YEAST
           547
                                   201
                                    ò
                                                           g
                                                                                    ò
                                                                                                          g
                                                                                                                                     δ
                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VLESDL----MQFKHISS------NEYII--EDSFKLLNSEQKNILLK-SYKYI 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              430 MLENELNNAALLLEHTSNEKNAKVKELNAKNOKLVECENDLOTLTKORLDLCROLOYLLI 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PSSPPTTPPSPAKTDEOKKESKFLPFLTNIETLYNNLVNKIDDYLINLK------AKIN 200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7 DNILSGFENEYDVIYLKPLAGVYRSLKKQIEKNIITFNLNLN---DILNSRLKKRKYFLD 63
                                                                                                                                                                                                                                                                                       Remacha M., Jimenez A., del Rey F., Ballesta J.P.G., Revuelta J.L.; The complete sequence of a 15,820 bp segment of Saccharomyces cerevisiae chromosome XI contains the UBIZ and MPLI genes and three new open reading frames.";
                                                                                                                                                                                                                                                                                                                                                                             ·i· FUNCTION: MYOSIN-LIKE PROTEIN THAT IS PROBABLY INVOLVED IN DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        105 KESVENDIK -- FAQEGISYYEKVLAKYKDDLESI-----KKVIKEE------KEKF
                                                                                                                                     Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
                                                                                                                                                                                                            MEDLINE-93247549; PubMed-8483450;
Koelling R., Nguyen T., Chen E.Y., Botstein D.;
"A new yeast gene with a chosin-like heptad repeat structure.";
Mol. Gen. Genet. 237:359-369(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1875;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COILED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
R -> A (IN REF. 1).
MW; 683A0D34C9066867 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY: SOME, TO THE TPR ONCOGENE.
CAUTION: REF.2 MISQUOTES THE GENE NAME AS "MPL1".
                                                                                                                                                Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6.5%; Score 132.5; DB 1; 21.6%; Pred. No. 8.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Coiled coil; DNA
                                                         01-OCT-1993 (Rel. 27, Created)
01-JUN-1994 (Rel. 29, Last sequence update)
115-JUL-1998 (Rel. 36, Last annotation update)
MLPI OR YKR095W OR YKR415.
                                   1875 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mismatches
                                                                                                                       Saccharomyces cerevisiae (Baker's yeast).
                                   PRT;
                                                                                                                                                                                                                                                                                      MEDLINE=94205265; PubMed=8154186;
Bou G., Esteban P.F., Baladron V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      pattern;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; L01992; AAA34783.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; X73541; CAA51948.1; -. EMBL; Z28320; CAA82174.1; -. PIR; S38173; S38173. SGD; SO001803; MLP1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   218455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1678
1866
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      301 30
1875 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
Matches 86; Conserv
                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Myosin; Heptad
                                   MLP1_YEAST
Q02455;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
          RESULT 13
MLP1_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        64
                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration
----AHVKITKLSDLKA-IDDKI-DL-----FKNTNDFEAI 238
                                                                                                                                                                                239 KKLIND - DTKKDM - - - LGK - - - - - LLSTGLVQIFPNTI - ISKLIEGKFQDMLNISQHQC 287
                                                                                                                                                                                                                                              related protein that
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kilmartin J.V., Dyos S.L., Kershaw D., Finch J.T.;
"A spacer protein in the Saccharomyces cerevisiae spindle poly body
whose transcript is cell cycle-regulated.";
J. Cell Biol. 123:1175-1184(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL
JECULES. IT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IS ESSENTIAL FOR GROWTH.
SUBCELLULAR LOCATION: NUCLEAR; TIGHTLY ASSOCIATED WITH THE
NUCLEUS. IT IS PRESENT IN A GRANULAR PATTERN THAT EXCLUDES THE
                                                                                          602 ELQKELEELKTSVPNEDASYSNVTIKQLTETKRDLESQVQDLQTRISQITRESTENMSLL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Favello A., Fulton L., Gattung S., Greco T., Kirsten J., Favello A., Fulton L., Gattung S., Greco T., Kirsten J., Fuchnson D., Johnston L., Langston Y., Latreille P., Le T., Johnston L., Langston Y., Latreille P., Le T., Mardis E., Meneces S., Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L., Taich A., Trevaskis E., Vignati D., Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.; Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEOSKELETON. POTENTIA. ROLE IN CROSSLINKING FILAMENTS OR ANCHORING OTHER MOLECULES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mirzayan C., Copeland C.S., Snyder M.; "The NUF1 gene encodes an essential coiled-coil related is a potential component of the yeast nucleoskeleton."; J. Cell Biol. 116:1319-1332(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-0CT-1993 (Rel. 27, Created)
01-0CT-1993 (Rel. 27, Last sequence update)
01-0CT-1996 (Rel. 34, Last annotation update)
NUFI PROTEIN (SPINDLE POLY BODY SPACER PROTEIN SPC110)
NUFI OR SPC110 OR YDR356W OR D9476.3.
Saccharomyces cerevisiae (Baker's yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Saccharomycetales; Saccharomycetaceae; Saccharomyces
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PTM: MAY BE REGULATED BY PHOSPHORYLATION EVENTS.
                                                                                                                                                                                                                                                                                                                                                                      288 VKKQCPENSGCFRHLDEREECKC-----LLNYKQE 317
                                                                                                                                                                                                                                                                                                                                                                                                                               944 AA
```

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
      엽.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ολ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Db
                                                                                                                                                                                                                                                                                                             18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                         84 ------DSF--KLLNSEQKNIL---LKSYKYIKESVENDIKFAQEGISYYEKVLAK 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         439 LQKKVKQLENDLFVIKKTHSESKTITDNELESKDKLIKILENDLKVAQEKYSKMEKELKE 498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                129 YKDDLESIKKVIKEEK----EKFPSSPPTTPPSPAKTDEQKKESKFLPFLTNIETLYNNL 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 185 VNKIDDYLINLK---AKINDCNVEKDEAHVKIT--KLSDLKAIDDKI-DLFKN----TND 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                235 FEAIKKLINDDTKKDMLGKLLSTGLVQIFPNTIISKLIEGKFQDMLNISQHQCVKKQCPE 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-S288C;
MEDLINE-92317149; PubMed-1618897;
HOYL M.A., He L., Loo K.K., Saunders W.S.;
"Two Saccharomyces cerevisiae kinesin-related gene products required
                                                                                                                                                                                                                                                                                                             Gaps
                                                                               protein; Phosphorylation; Heptad repeat pattern
                                                                                                                                                                                                                                                                                                                                                  27 GVYRSLKKQIEKNIITFNLNLNDI-LNSRLKKRKYFLDVLESD--LMQFKHISSNEYIIE 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              295 NSGCFRHLDEREECKCLLNYKQEGDKCEENPNPTCNENNGGCDADATCTEEDSGSSRKK 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----IESEDELKSL----RDSQKIEIE-NWKRKYNNLSLENDRLLTEKESASDKER 703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R., Aviles E., Bernon A., Brennan T., Carpenter J., Chenr B., Cherry J.M. Chung E., Duncan M., Guzman E., Hartzell G., Hunicke-Smith S., Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Ocfner P., Oh C., Petel F.X., Roberts D., Schl P., Schramm S., Shogren T., Smith V., Taylor P., Wel Y., Yelton M., Botstein D., Davis R.W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Saccharomyces cereiisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                             . 89
                                                                                                                                                                                                                                                                     DB 1; Length 944; 4.4;
                                                                                                                                                                                                                                                                                                         64; Mismatches 142; Indels
                                                                                                                                                                                                         04FAA074BB8A0BC8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-AUG-1992 (Rel. 23, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
CINB OR KSL2 OR YEL061C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ą
                                                                                                      COILED COIL.
                                                                                                                                                                                   ARG/TYR-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 1038
                                                                                                                                                                                                                                                                  6.4%; Score 132; 23.7%; Pred. No. 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               for mitotic spindle assembly."; J. Cell Biol. 118:109-120(1992)
                                                                                                                                                                                                       111781 MW;
                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 23.78
Matches 85; Conservative
EMBL; U28372; AAB64791.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                     59
731
747
944
                   PIR; S26710; S26710.
PIR; S34288; S34288.
SGD; S0002764; NUF1.
                                                                             Coiled coil; Nuclear
DOMAIN 164 79
                                                                                                                                                                                                    944 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CIN8_YEAST
                                                                                                                                                                                   DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            P27895;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 15
CIN8_YEAST
1D CIN8_YEAST
1D CIN8_YEAST
1D CIN8_YEAST
DT 01-AUC
DT 01-AUC
DT 01-AUC
DT 05-YEE
DT 01-AUC
OS SACCH
OS SACC
                                                                                                                                                                DOMAIN
                                                                                                                       DOMAIN
                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      657
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRINTS; PR00380; KINESINHEAVY.
PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
MOCOF protein; Cell division; Microtubules; ATP-binding; Coiled coil; Mitosis; Cell cycle.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CIN8 AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      87 KLLNSEQKNILLKSYKYIKESVENDIKFAQEGISYYEKVLAKYKDDLESIKKVIKEEKEK 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Rousselet G., Simon M., Ripoche P., Buhler J.M.;
"A second nitrogen permease regulator in Saccharomyces cerevisiae.";
FEBS Lett. 359:215-219(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                           mitotic spindle.";
Cell 70:451-458(1992).
-!- FUNCTION: REQUIRED FOR ASSEMBLY OF THE MITOTIC SPINDLE. INTERACT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WITH SPINDLE MICROTUBULES TO PRODUCE AN OUTWARDLY DIRECTED FORCE ACTING UPON THE POLES. FOLLOWING SPINDLE ASSEMBLY, CINB ANK KIPL APPARENTLY ACT TO OPPOSE A FORCE THAT DRAWS SEPARATED POLES BACK TGGETHER, THIS FORCE SEEMS TO BE MEDIATE BY KAR3. SUBCELLULAR LOCATION: SPINDLE MICROTUBULES THAT LIE BETWEEN THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               53; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29 YRSLKKQIEKNIITFNLNLNDILNSRLKKRKYFLDVLESDLMQFKHISSNE--YIIEDSF 86
                                                                                                                                                                                                                                                                                                                                                                                Saunders W.S., Hoyt M.A.; "Kinesin-related proteins required for structural integrity of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              147 FPSSPPTTPPSPAKTDEQKKESKFLPFLTNIE-TLYNNLVNKIDDY----LINLKAKIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIMC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 1038;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MECHANOCHEMICAL (MOTOR) (BY SI
COLLED COLL (POTEWTIAL).
COLLED COLL (POTEWTIAL).
ATP (BY SIMILARITY).
D -> A (IN REF. 1).
Q -> H (IN REF. 1).
WW; 3AIPD7003EF89FBC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6.4%; Score 130.5; DB 1; 25.3%; Pred. No. 5.7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54; Mismatches 126;
                                                                                         STRAIN=S288C;
MEDLINE=95172238; PubMed=7867803;
                                                                                                                                                                                                                                                                                                                                                  MEDLINE=92354062; PubMed=1643659;
                                                        SEQUENCE OF 1021-1038 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; 211859; CAA77885.1; -. EMBL; M90522; AAA34496.1; -. EMBL; U18795; AAB65026.1; -. EMBL; X79105; CAA55722.1; -. PIR; B42641; B42641;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                117999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001752; -.
Pfam; PF00225; kinesin; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HSSP; P17119; 3KAR. SGD; S0000787; CIN8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
Matches 79; Conserv
                                                                                                                                                                                                                                                                                   CHARACTERIZATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUBFAMILY.
                                                                                                                                                                                                                                                                                                                 STRAIN-S288C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POLES
```

15;

```
qq
ò
```

	902	
•	HITKEIKKLKSTLFLQLNTMQQESILQETNIQPNLDMI	
	MODIN	
	)ET)	
	ĭ	
	2	
	<b>LFLOLNTMQQESILQETNIQ</b>	
-	LINI	
-	LFI.	
	KST	
•	꿏	
	EIK	
-	ITK	
-	H	
-	臽	
	Ϋ́F	
	0	
-	662 LN-QKFEMH-	
	662	

ò

Search completed: August 8, 2001, 12:32:34 Job time: 202 sec

QQ

<sup>322</sup> EENPNPTCNENN 333 Q D

<sup>|| : |||</sup> || 362 AENLSDISEENN 773